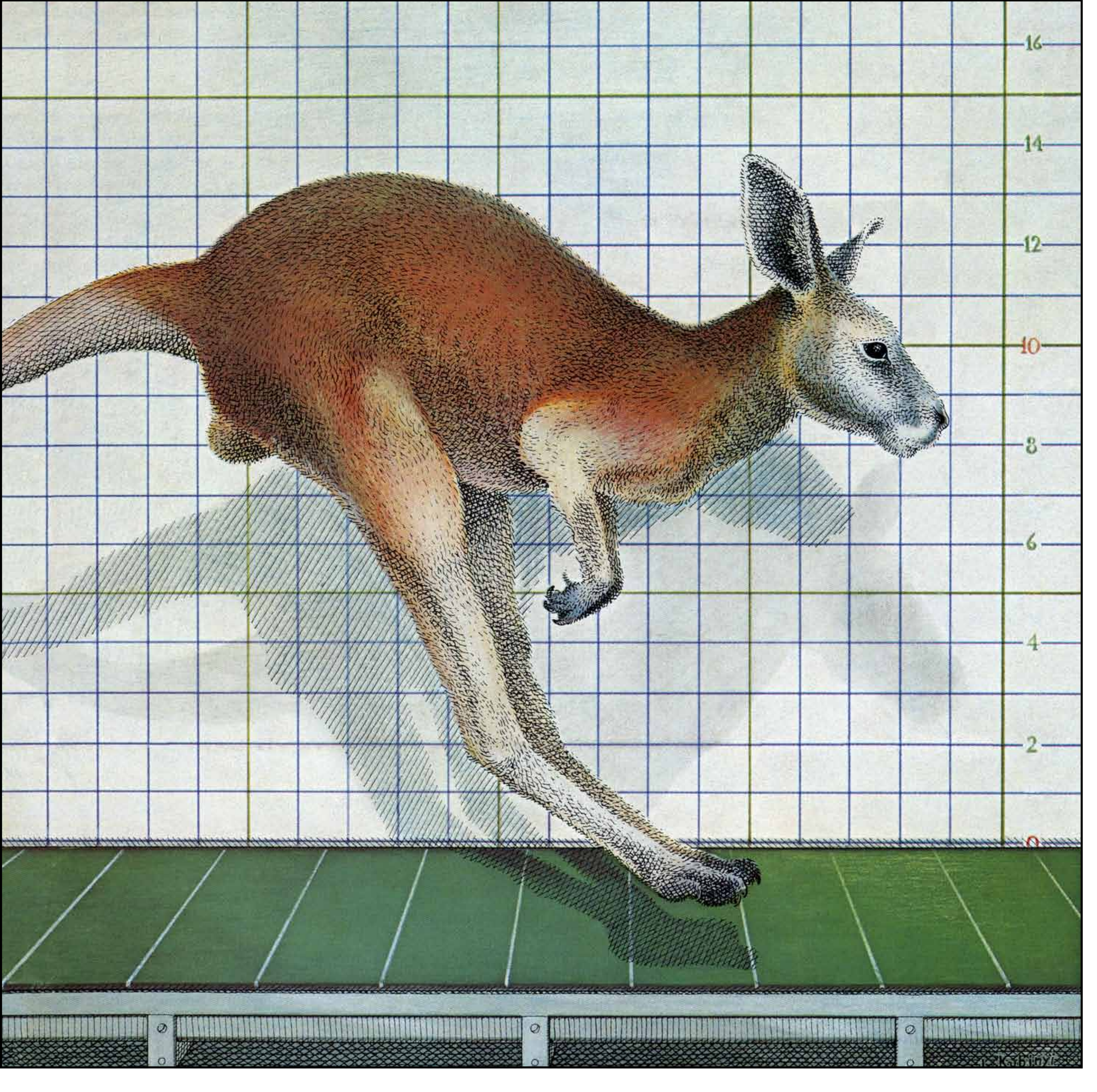
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



CANGUROS

Octubre 1977

140 Ptas.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 13
Octubre 1977

INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN

4 **LAS NEGOCIACIONES SALT, Herbert Scoville, Jr.**
Las conversaciones tratarán de ambas carreras armamentísticas: cuantitativa y cualitativa.

14 **OBJETOS BL LACERTAE, Michael J. Disney y Philippe Véron**
Con los quasars, estos objetos se cuentan entre los más remotos y más luminosos del universo.

24 **COMUNICACIONES OPTICAS, W. S. Boyle**
El primer ensayo comercial con servicio telefónico óptico se encuentra ya funcionando.

38 **FLUJO DE CALOR DESDE EL INTERIOR DE LA TIERRA, Henry N. Pollack y David S. Chapman.** El modelo resultante se interpreta en términos de la tectónica de placas.

52 **CANGUROS, T. J. Dawson**
Su adaptación a la vida en las sabanas de un continente aislado es realmente única y extraordinaria.

62 **LOS VACUOLOS GASIFEROS DE LAS CIANOFICEAS, A. E. Walsby**
Las células de muchas especies acuáticas contienen miles de estructuras diminutas, llenas de gas.

72 **PERSPECTIVAS EN QUIMICA MACROMOLECULAR, Juan A. Subirana**
La hipótesis macromolecular ha descifrado la naturaleza de los plásticos y los secretos de la vida.

82 **"SEGUNDOS MENSAJEROS" EN EL CEREBRO, James A. Nathanson y Paul Greengard**
Ciertos mensajes químicos entre células nerviosas se traducen en una serie de efectos fisiológicos.

3 AUTORES

34 CIENCIA Y SOCIEDAD

96 JUEGOS MATEMATICOS

102 TALLER Y LABORATORIO

108 LIBROS

112 BIBLIOGRAFIA

EDITA Prensa Científica, S.A.
Calabria, 235-239
Barcelona-15
ESPAÑA

Colaboradores de este número:

Asesoramiento y traducción:

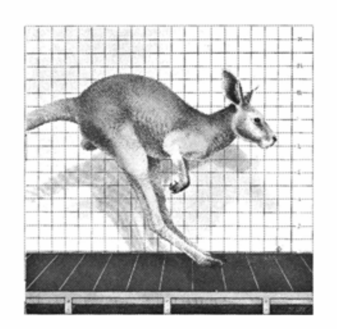
Francisco L. de Sepúlveda: *Las negociaciones SALT;* Ricardo Vieira: *Objetos BL Lacertae;* Jesús Díaz Cortijo (C. Inves. CTNE): *Comunicaciones ópticas;* Miquel Gich: *Flujo de calor desde el interior de la tierra;* Carles Enric Riba: *Canguros;* Marta Estrada: *Los vacúolos gasíferos de las cianofíceas;* Monique Robert: *"Segundos mensajeros" en el cerebro;* José M.ª Vidal: *Taller y laboratorio;* Luis Bou: *Juegos matemáticos.*

Ciencia y sociedad:

Manuel García Doncel

Libros:

Luis Alonso, Antonio Comella y Philip Morrison



LA PORTADA

La ilustración de la portada representa un experimento de laboratorio sobre la dinámica de la energía animal, llevado a cabo en la Universidad de Harvard. Se trataba de comparar el gasto energético en la marcha a saltos de dos patas con la marcha de las cuatro extremidades. De los grandes mamíferos, los únicos que se mueven a saltos son los canguros australianos. Se llevaron cuatro canguros rojos desde New South Wales hasta Massachusetts, con la intención de medir sus saltos en una cinta corredera. De los cuatro, dos llegaron a saltar tal y como muestra la ilustración. A lo largo del experimento se colocaban máscaras a los canguros que saltaban en la cinta, para medir cuánto oxígeno consumían y el gasto energético del salto.

Suscripciones:

Prensa Científica, S.A.
Departamento de suscripciones
Calabria, 235-239
Barcelona-15 (España)

Distribución para España:

Distribuciones de Enlace, S.A.
Ausias March, 49
Barcelona-10 (España)

Distribución para los restantes países:

Editorial Labor, S.A.
Calabria, 235-239
Barcelona-15 (España)

Precio de venta de este número:

España: 140 pesetas
Extranjero: 2 U.S. $

Condiciones de suscripción:

España:
Un año (12 números): 1.540 pesetas
Extranjero:
Un año (12 números): 30 U.S. $

Ejemplar atrasado ordinario:
160 pesetas

Ejemplar atrasado extraordinario:
225 pesetas.

Solicitado control **OJD**

PROCEDENCIA DE LAS ILUSTRACIONES

Ilustración de la cubierta de Laszlo Kubinyi

Página	Fuente
5	U. S. Air Force *(arriba); Jane's All the World's Aircraft (abajo)*
6-11	Gabor Kiss
15	Wise Observatory, Tel-Aviv University *(arriba, izquierda y centro);* Observatorio de Haute-Provence *(arriba, derecha);* National Geographic Society-Palomar Observatory Sky Survey *(abajo)*
16-21	Adolph E. Brotman
25	Ralph Morse
26-32	Allen Beechel
35	CERN
36	Nicola Zanichelli
39	Wendell A. Duffield, U. S. Geological Survey
40	Henry N. Pollack y David S. Chapman
41-50	Andrew Tomko
53	Lorelle M. Raboni
54	R. McBlain
55-60	Lorelle M. Raboni
63	Daniel Branton, Harvard University
64	A. E. Walsby, University College of North Wales
65-68	Ilil Arbel
70	Daniel Branton, Harvard University
73	J. A. Subirana y A. Sellés
74	J. Kendrew *(arriba);* J. A. Subirana *(abajo)*
75	M. Pladellorens y J. A. Subirana *(arriba);* F. Azorín y J. A. Subirana *(abajo)*
76	J. A. Subirana y A. Sellés *(arriba);* S. Muñoz y J. A. Subirana *(abajo)*
77	J. A. Subirana
78	J. A. Subirana y A. Sellés *(arriba);* J. A. Subirana *(abajo)*
79	J. A. Subirana y A. Sellés
80	J. A. Subirana *(arriba);* SCIENTIFIC AMERICAN *(abajo)*
83	Jerome Kuhl
84	Noel T. Florendo, Baptist Memorial Hospital, Memphis, Tenn.
85-87	Jerome Kuhl
89	Alan D. Iselin
90	Jerome Kuhl
91	Bruce K. Krueger, Washington University
92-94	Jerome Kuhl
96-100	Ilil Arbel

Dep. legal: B. 38.999-76
Fotocomposición: Fort, S.A.
Rosellón, 33 - Barcelona-15
Cayfosa. Santa Perpetua de Moguda. Barcelona.

Printed in Spain - Impreso en España

Los autores

HERBERT SCOVILLE, JR. ("Las negociaciones SALT") es secretario de la Arms Control Association, que es un organismo privado con sede en Washington. Después de graduarse en la Universidad de Yale en 1937, realizó trabajos de investigación en química física en las de Cambridge y Rochester, obteniendo el doctorado por esta última en 1942. Scoville preside hoy el Council on Foreign Relations SALT Working Group.

MICHAEL J. DISNEY y PHILIPPE VÉRON ("Objetos BL Lacertae") trabajan en observatorios astronómicos. Disney se doctoró en astronomía teórica por la Universidad de Londres en 1968. Sin embargo, tras dos años de docencia en la Universidad de Arizona, prefirió la vida nómada del astrónomo observacional, pasando cuatro años en el observatorio Mount Stromlo en Australia y otros dos en la Universidad de Groningen, en Holanda. Hoy se encuentra en el Royal Greenwich Observatory, en Sussex. Véron, de nacionalidad francesa, tras su paso por la École Polytechnique parisiense, empezó, en 1962, a estudiar radioastronomía en el Observatorio de París, y después de trabajar durante un año con Allan R. Sandage en los Hale Observatories de Pasadena, se doctoró en 1967. En la actualidad disfruta de dos años de excedencia del European Southern Observatory de Ginebra.

W. S. BOYLE ("Comunicaciones ópticas") es director de la Research and Communications Sciences Division, de los Bell Laboratories. Nacido en Nova Scotia, pilotó Spitfires de la Royal Navy durante la Segunda Guerra Mundial. Finalizada la cual, estudió en la Universidad McGill, obteniendo su doctorado en física en 1950. Se incorporó a los Bell Laboratories en 1953 y participó en el desarrollo de láseres de estado sólido. Más tarde, con Donald F. Nelson, de los Bell Laboratories, desarrolló el primer láser de ruby de operación continua. Boyle es coinventor del dispositivo de carga acoplada, una técnica microelectrónica de grandes aplicaciones.

HENRY N. POLLACK y DAVID S. CHAPMAN ("Flujo de calor desde el interior de la tierra") son, respectivamente, profesor de geofísica en la Universidad de Michigan y profesor adjunto de geofísica en la Universidad de Utah. Pollack estudió geología en la Universidad de Cornell y en la de Nebraska y se doctoró en geofísica por la Universidad de Michigan en 1963. Posteriormente, realizó su investigación postdoctoral con el grupo de geotermia dirigido por Francis Birch, de la Universidad de Harvard. En 1970, Pollack conoció a Chapman mientras trabajaba en mediciones geotérmicas en los prominentes altiplanos de Africa Central. Chapman es de origen canadiense y estudió en la Universidad de British Columbia. Durante seis años permaneció como profesor voluntario en Zambia, primero en una escuela secundaria y posteriormente impartiendo clases de geofísica en la Universidad Nacional.

T. J. DAWSON ("Canguros") es profesor y director de la facultad de zoología de la Universidad de New South Wales, Australia. "Mi interés por la biología de los animales australianos se remonta a mi infancia, cuando merodeaba por entre los matorrales de los alrededores de la pequeña ciudad de Nyngan, en el interior de la región lanera de Australia." Después de doctorarse en la Universidad de New England, Australia, en 1964, estuvo tres años en los Estados Unidos, inicialmente con Knut Schmidt-Nielsen en la Duke University y posteriormente con Theodore Hammel en la de Yale. Desde su regreso a Australia ha venido trabajando en la fisiología y ecología de los marsupiales y monotremas, con especial énfasis en el canguro.

A. E. WALSBY ("Los vacúolos gasíferos de las cianofíceas") es lector en el departamento de biología marina en el University College de North Wales. Tras graduarse en botánica por la Universidad de Birmingham, investigó, en unión con G. E. Fogg, las cianofíceas en la Universidad de Londres, en donde se doctoró en 1965. De 1971 a 1973 fue miembro de la Universidad de California en Berkeley: allí continuó sus trabajos sobre tales algas en colaboración con Daniel Branton.

JUAN ANTONIO SUBIRANA ("Perspectivas en química macromolecular") es catedrático de la Escuela Técnica Superior de Ingenieros Industriales de Barcelona y director del departamento de química macromolecular del C.S.I.C. Estudió química e ingeniería industrial en Barcelona y Madrid, doctorándose en ambas especialidades. Posteriormente trabajó como investigador en las Universidades de París y Harvard y en el Instituto Weizmann de Israel. Está afiliado a numerosas sociedades científicas y es miembro del Comité de expertos en biofísica de la región europea de la UNESCO y de la Organización Europea de Biología Molecular. En el departamento que dirige se realizan actualmente trabajos de investigación en distintas áreas de la química macromolecular, entre las que cabe destacar: fisicoquímica de polielectrolitos y de su comportamiento reológico; caracterización de polímeros sintéticos por microscopía electrónica y difracción de rayos X; estructura y composición de los cromosomas; proteínas de los espermatozoides y su interacción con el ácido desoxirribonucleico; espermiogénesis humana y sus anomalías.

JAMES A. NATHANSON y PAUL GREENGARD ("'Segundos mensajeros' en el cerebro") están ambos en la facultad de medicina de la Universidad de Yale. Nathanson ha mostrado interés por la ciencia desde la escuela primaria, cuando realizaba experimentos sobre aprendizaje y memoria en un laboratorio que tenía montado en el sótano de su casa. Fue al Trinity College y luego estudió medicina en Yale. Durante su carrera en Yale, se preocupó por el enfoque bioquímico de la función cerebral y, siguiendo con esta línea, se doctoró en neurobiología en el laboratorio de Greengard. Después de graduarse, estuvo un año como médico interno, pero luego pasó otros dos en los National Institutes of Health. En 1976, volvió a Yale como residente en neurología. Greengard es profesor de farmacología en Yale. La ciencia le interesó por primera vez cuando, a la edad de 17 años, estando alistado en la Marina durante la segunda guerra mundial, le mandaron al Instituto de Tecnología de Massachusetts para que participara en la elaboración de un sistema de alarma precoz, destinado a proteger a los navíos de la Marina del ataque de los aviones suicidas japoneses. Después de la guerra, fue al Hamilton College y luego a la Universidad de Johns Hopkins, donde la rama objeto de su interés evolucionó de las matemáticas y la física a la neurobiología. Tras doctorarse en Johns Hopkins en 1953, estuvo cinco años realizando trabajos postdoctorales en Inglaterra.

Las negociaciones SALT

Las conversaciones en curso sobre limitación de armamentos estratégicos pueden poner un techo efectivo a la carrera cuantitativa de armamentos y constreñir la carrera cualitativa

Herbert Scoville, Jr.

La Administración del presidente Carter tiene sobre la mesa la revisión de la política de Seguridad Nacional y la preparación de su primer presupuesto de Defensa. Una cuestión importante a dilucidar se centra en el papel que puede y debe desempeñar el control de armamentos, así como las conversaciones sobre limitación de armamentos estratégicos (SALT) con la URSS, que son las negociaciones más difíciles actualmente en curso. Tanto la Unión Soviética, con sus nuevos modelos de misiles balísticos intercontinentales (ICBM) de gran capacidad de carga, como los Estados Unidos, con sus programas para reemplazar sus ICBM Minuteman por otros de mayor precisión, tienen a la vista nuevos proyectos que podrían comprometer de forma considerable la estabilidad del equilibrio estratégico, además de costar miles de millones de dólares. El acuerdo provisional sobre armas ofensivas logrado en 1972 expira el día 3 de este mes y, a menos que se llegue a un acuerdo con anterioridad a esa fecha, las presiones para acelerar la carrera de armamentos serán muy grandes. El éxito de esas conversaciones es, por tanto, un asunto de seria trascendencia tanto para la seguridad de los Estados Unidos como para la cuenta de gastos de armamentos que tendrán que pagar los norteamericanos.

Los Estados Unidos y la Unión Soviética llegaron a un primer acuerdo sobre el control de armamentos en 1972, cuando se firmó el tratado sobre misiles antimisiles balísticos y el acuerdo provisional, al finalizar la primera tanda de conversaciones sobre limitación de armamentos estratégicos (SALT I). Las dos superpotencias están inmersas, desde entonces, en la segunda ronda de las conversaciones (SALT II). El plazo de validez del acuerdo provisional está expirando ya que se estableció tan sólo para cinco años. Hace casi tres años se dio un paso importante con vistas a remontar la fecha de expiración que vence ahora. Los acuerdos de Vladivostok fijaron unas líneas de acción con respecto a las SALT II, estableciendo los techos siguientes para las principales armas estratégicas: 2400 vectores estratégicos (entre bombarderos y misiles), de los cuales únicamente 1320 podrían ir provistos con ojiva de múltiples cabezas (MIRV) [MIRV corresponde a la sigla de "multiple independently-targetable reentry vehicle" (vehículo de reentrada con objetivos múltiples asignables independientemente) que designa, en un misil ICBM o SLBM, un tipo de ojiva en cuyo interior van varias cabezas nucleares que se desprenden de la misma y se dirigen contra objetivos diferentes una vez el misil vuelve a entrar en la atmósfera, ya sobre territorio enemigo. N. del T.]. La parte positiva de dichos acuerdos fue que, por vez primera, se establecía un techo en cuanto al número total de bombarderos y de misiles, así como en la cifra de misiles que podían armarse con MIRV (aunque sin determinar el número total de cabezas independientes). En el aspecto negativo, cabe decir que los techos fijados eran tan elevados que prácticamente no suponían limitación alguna. También hay que destacar la imprecisión en lo que respecta a la definición de qué tipos de armas –en particular misiles crucero y bombarderos de alcance medio– quedarían incluidas dentro de los techos fijados y, por otra parte, que al no concretar nada acerca de los avances tecnológicos, se institucionalizaba, de hecho, una nueva carrera de armamentos: la cualitativa, en contraposición a la cuantitativa. Todo esto reviste importancia, como se ha demostrado posteriormente. Lo que complica el desarrollo actual de las negociaciones SALT es que cada parte está intentando jugar –de modo diferente– con las inconcreciones derivadas de los acuerdos de Vladivostok.

La nueva Administración del presidente Carter tomó la iniciativa en marzo pasado con ocasión del viaje del secretario de estado, Cyrus R. Vance, a Moscú para presentar dos propuestas alternativas. La primera de ellas era "comprehensiva". Por un lado, trataba de frenar la carrera cuantitativa reduciendo los arsenales existentes realmente y, por otro, frenar la carrera cualitativa restringiendo la modernización o sustitución de los sistemas de armas ahora desplegados. La segunda propuesta era una proposición "expectante". Consistía simplemente en confirmar los techos de Vladivostok, dejando para futura consideración los dos elementos que bloquearon la consecución de unos acuerdos SALT II a lo largo del mandato del presidente Ford: el misil crucero norteamericano y el bombardero medio Backfire soviético. Los rusos rechazaron ambas propuestas. La mencionada en segundo lugar era, aparentemente, satisfactoria para los soviéticos; pero no podían aceptar un aplazamiento en cuanto al misil crucero. Con respecto a la primera propuesta, los soviéticos dijeron que las limitaciones y prohibiciones que contenía estaban calculadas con vistas a congelar la superioridad estadounidense.

Después de esta negativa inicial, las negociaciones prosiguieron a partir de un acuerdo de trabajo sobre tres bases concretas, establecido en el mes de mayo en Ginebra. Se llegaría, primeramente, a un tratamiento de cinco años de duración confirmando los techos de Vladivostok o tal vez otros techos inferiores; en segundo lugar, a un protocolo de tres años de validez, limitando el desarrollo de misiles crucero, misiles móviles, ICBM pesados y el bombardero Backfire y, en tercer lugar, a una declaración de principios para la continuación de negociaciones a largo plazo, reducciones sustanciales en armamentos estratégicos y acerca del control de la carrera cualitativa. Las negociaciones desarrolladas durante este verano han versado, presumiblemente, sobre los detalles a introducir en este acuerdo de trabajo.

A continuación voy a describir y analizar, principalmente, los detalles de la primera de las propuestas presentadas por los Estados Unidos en Moscú, intentando al propio tiempo sintetizar la problemática de los aspectos fundamentales del control de armamentos, las complejas interrelaciones entre los controles cualitativo y cuantitativo y las dificultades para llegar a controles eficaces en una era en la que los avances tecnológicos son mucho más importantes que la medición convencional de la supremacía estratégica basada en el número y la capacidad en megatones de los misiles de cada bando.

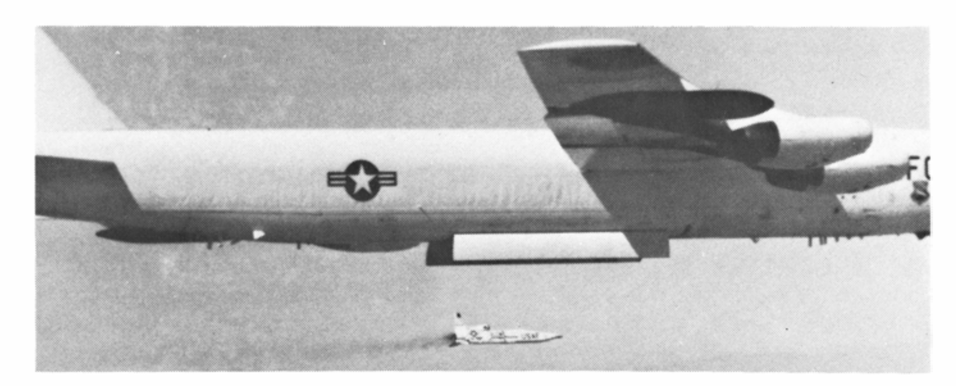
La propuesta norteamericana "comprehensiva" iba mucho más allá de cualquier tipo de control que se hayan planteado mutuamente las dos potencias a lo largo de ocho años de negociaciones. Suponía, especialmente, un esfuerzo por limitar los sistemas de armas con mayores posibilidades para poner en peligro el equilibrio de disuasión mutua entre las dos potencias, es decir, sistemas que puedan dar a cualquiera de ellas una capacidad de "contrafuerza" ("counterforce") al poner en peligro los misiles balísticos intercontinentales con base en tierra de la otra. Un arma de "contrafuerza", por ejemplo un ICBM armado con MIRV, es la que cuenta con muchas probabilidades de destruir una base de lanzamiento –un asentamiento en silo subterráneo protegido– combinando la alta precisión con la potencia. Pone en peligro a un elemento esencial del arsenal de represalia del otro bando y, por tanto, a su fuerza de disuasión.

Los nueve puntos básicos de la propuesta norteamericana (según han sido dados a conocer hasta la fecha) se agrupaban en dos categorías. Los cuatro primeros establecían límites –más bajos que los actualmente en vigor– en el número de vectores estratégicos. Los cinco últimos abordaban el problema del control cualitativo imponiendo varias restricciones al desarrollo, pruebas y despliegue de nuevas armas.

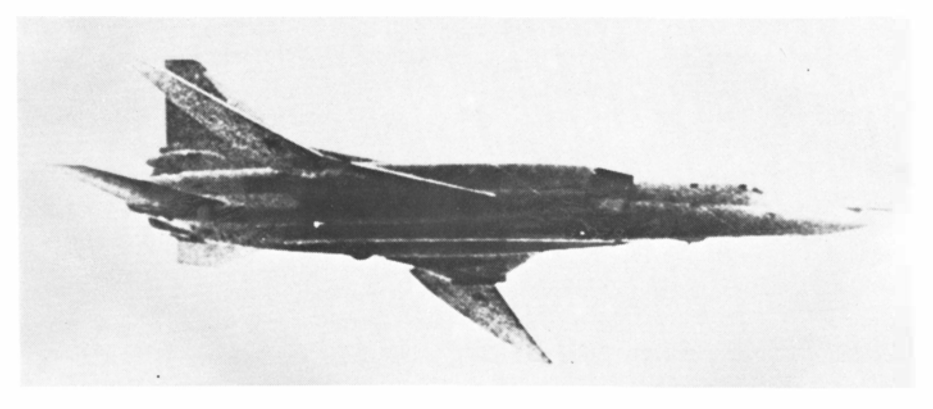
Las medidas cuantitativas reducirían el número total de vectores, es decir, de ICBM, misiles balísticos lanzados desde submarino (SLBM) y bombarderos estratégicos, pasando del total de 2400 acordado en Vladivostok a una cifra comprendida entre 1800 y 2000; de ellos, solamente podrían ser armados con MIRV entre 1100 y 1200 (en contraposición con los 1320 establecidos en Vladivostok). Unicamente 550 de los misiles con MIRV podrían ser ICBM (una condición parcial sobre la que nada se acordó en Vladivostok). Por último, el número de ICBM pesados soviéticos, fijado en 308 según el acuerdo provisional, quedaría reducido a la cifra de 150.

Si atendemos al modo en que cada bando podría distribuir el total de vectores estratégicos permitidos –entre ICBM, SLBM y bombarderos–, a la potencia de las cabezas nucleares y al número de MIRV a colocar en cada tipo de misiles, salta a la vista la trascendencia que la disminución de los techos tendría en los arsenales de las dos potencias *(véanse las ilustraciones de las páginas 8 y 9)*. La distribución real podría modificarse, pero ello no alteraría probablemente el equilibrio estratégico en un grado significativo. (La mayor incertidumbre se centra en el programa soviético de equipar con MIRV los SLBM, un programa que hasta el momento no ha sido objeto de prioridad en Rusia, pero que presumiblemente cobraría importancia, como explicaré más tarde, si se redujera la cifra de ICBM dotados con MIRV.)

Los esfuerzos de la Unión Soviética se han concentrado en construir ICBM más potentes, capacitados para transportar mayor carga, lo cual permite, o bien mayor potencia explosiva por cabeza nuclear, o mayor número de cabezas por misil, o una solución intermedia entre las dos. Basándome en las pruebas realizadas por los soviéticos, opino que sus ICBM pesados podrían llevar ocho cabezas nucleares cada uno, y seis los ICBM más ligeros. Los ICBM norteamericanos Minuteman III van armados con tres cabezas. Con respecto a los bombarderos, supongo que los B-52 estadounidenses podrían transportar hasta 10 armas, bien sean del tipo bomba, misiles crucero de largo alcance o misiles aire-tierra de corto alcance. No hay datos seguros acerca de la capacidad de carga de los bombarderos soviéticos, pero el Departamento de Defensa estadounidense la estima en dos o tres armas por aparato: como la fuerza soviética de bombarderos estratégicos es muy pequeña, la trascendencia de esta falta de datos no es importante. En este cálculo he excluido al bombardero medio Backfire, ya que sus posibilidades en misiones intercontinentales son realmente discutibles. Puede asimilarse a los bombarderos estadounidenses estacionados en Europa, con ra-



EL MISIL CRUCERO NORTEAMERICANO es un reactor sin piloto, de vuelo subsónico, similar a la "bomba volante" de la Segunda Guerra Mundial, pero con mucho mayor alcance y precisión. Puede ser lanzado desde tierra, navío de superficie, submarino o desde el aire (como desde este B-52). Uno de los puntos básicos de las negociaciones SALT es la insistencia soviética en que los misiles crucero de largo alcance queden incluidos en los techos de armas estratégicas.



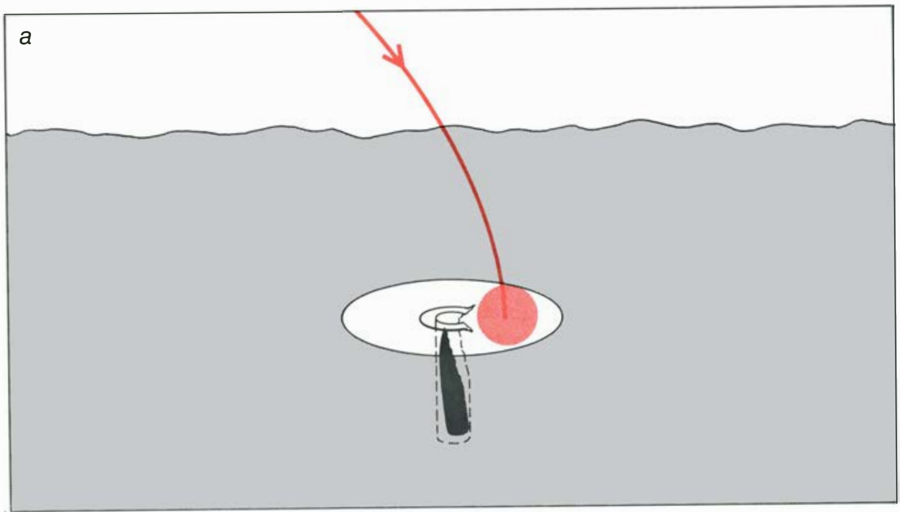
EL BACKFIRE SOVIETICO es un bombardero supersónico capaz de alcanzar una velocidad Mach 2 a gran altitud. Su radio de combate es de 5500 kilómetros, lo que significa que podría penetrar en los Estados Unidos a velocidad inferior a la del sonido desde las bases soviéticas instaladas en el Artico en una misión "suicida" (sin regreso), tomando tierra en un "tercer país" o reabasteciéndose en vuelo. La cuestión estriba en determinar si el Backfire es o no un arma estratégica.

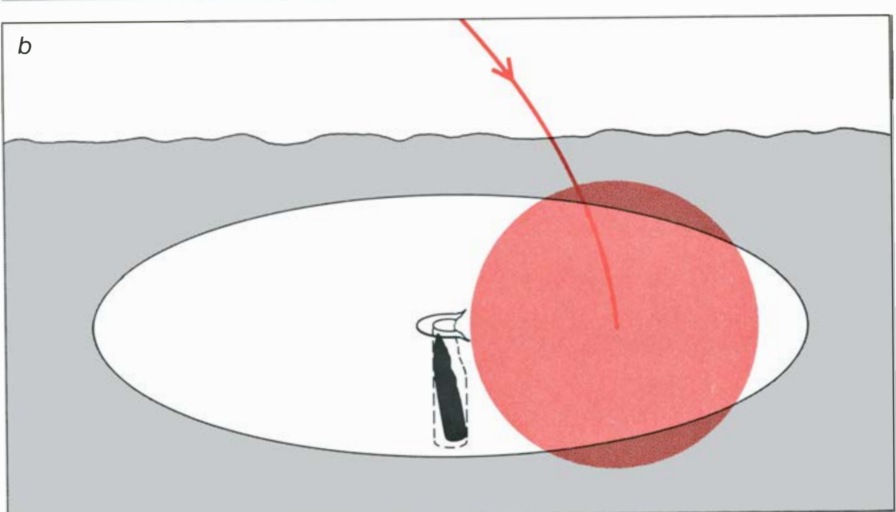
dio de acción sobre la Unión Soviética, que forman parte del denominado sistema de bases adelantadas (forward-based system).

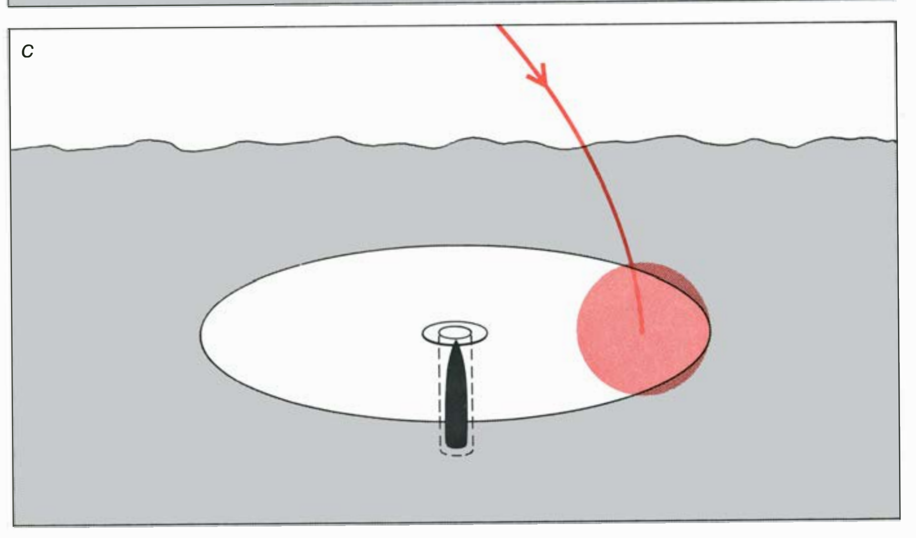
La adopción de los más bajos niveles de armamentos que contiene la propuesta "comprehensiva" presentaría varios tipos de problemas –tanto reales como basados en supuesta mala fe– para los Estados Unidos y la Unión Soviética. Dicha propuesta exigiría, por parte de los soviéticos, una reducción en el número de ICBM pesados de 308 a 150, todos los cuales podrían llevar cabezas nucleares múltiples. Esa condición provocó una fuerte reacción rusa, ya que como los Estados Unidos no tienen ICBM pesados, los términos de la propuesta no le afectaban. Este aspecto polémico de la cuestión podría haberse evitado si la propuesta se hubiese expresado en términos diferentes, simplemente estableciendo un techo de 150 en cuanto al número de misiles pesados que los rusos podrían dotar con MIRV, o incluso especificando el número de SS-9 soviéticos a reemplazar por los nuevos SS-18. Los SS-9 nunca se han experimentado con MIRV, y la mayoría de los SS-18 llevan una cabeza nuclear, por lo que los soviéticos sólo disponen en la actualidad de un reducido número de ICBM pesados con MIRV. Los techos en cuanto a ICBM dotados con MIRV o, en definitiva, el número de SS-18, podrían ser verificables mediante satélites, ya que los SS-18 requieren un silo subterráneo mayor que el de los SS-9. El establecimiento de unos techos sobre esta base tendría un efecto estratégico similar al de una reducción general de misiles, debido a que los SS-9 de cabeza nuclear sencilla no suponen una amenaza significativa a efectos de "contrafuerza".

La limitación del número de ICBM con MIRV a 550 sería particularmente difícil de aceptar por los soviéticos, ya que ellos conceden más importancia que los Estados Unidos a los ICBM (el límite de 550 presenta también un problema de "enmascaramiento", al haber sido escogido coincidiendo evidentemente con el número actual de misiles Minuteman norteamericanos dotados con MIRV). No puede saberse con seguridad cuántos de los 1320 vectores con MIRV acordados en Vladivostok podrían ser adjudicados por los soviéticos a su arsenal de ICBM, pero, en vista de la escasa prioridad que en el pasado concedieron a los SLBM, la cifra autorizada podría haber sido superior a 550. Por otra parte, como los soviéticos disponen hoy de menos de 200 ICBM con MIRV, esta condición no les hubiera obligado a desmantelar ninguno. También hay que tener en cuenta que si el techo en ICBM con MIRV se situase muy por encima de 550, la capacidad de "contrafuerza" sería una realidad a considerar, habida cuenta de los previsibles avances en punto a precisión.

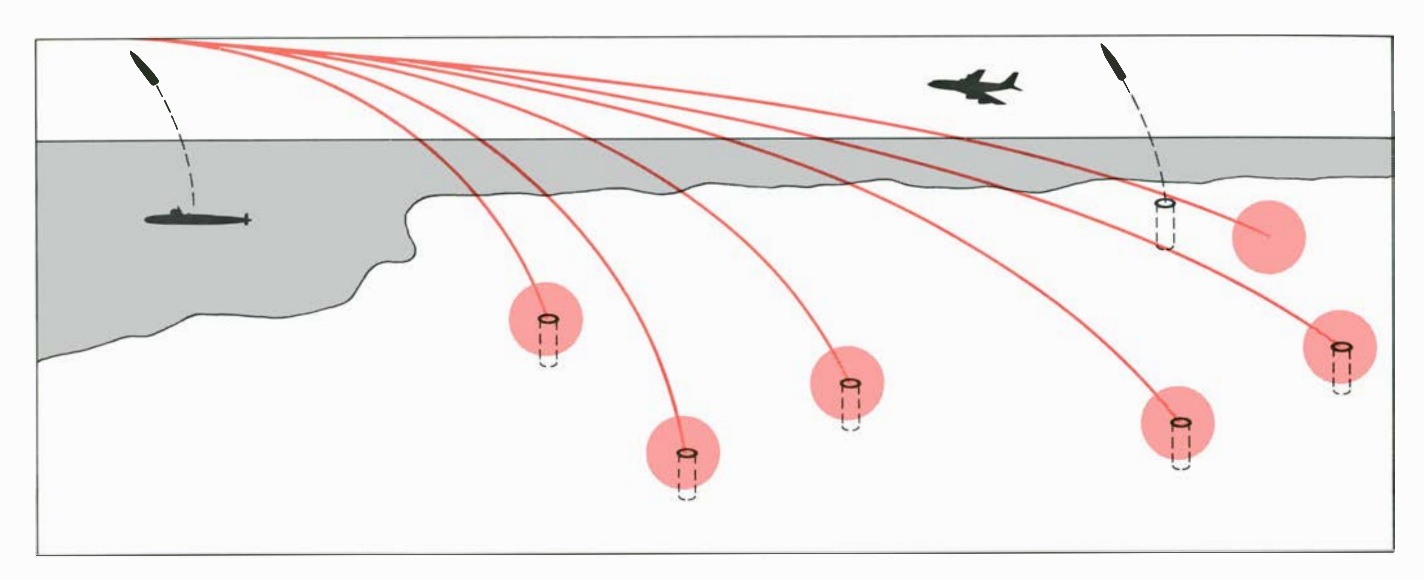
El simple hecho de establecer un techo de 1800 a 2000 vectores de lanzamiento obligaría probablemente a los rusos a retirar del servicio cierto número de ICBM. También tendrían la alternativa de reducir la cifra de SLBM, pero ésta ascendía en la primavera pasada a 840 unidades y parecer hallarse en camino de alcanzar el techo de los 950 permitidos por el



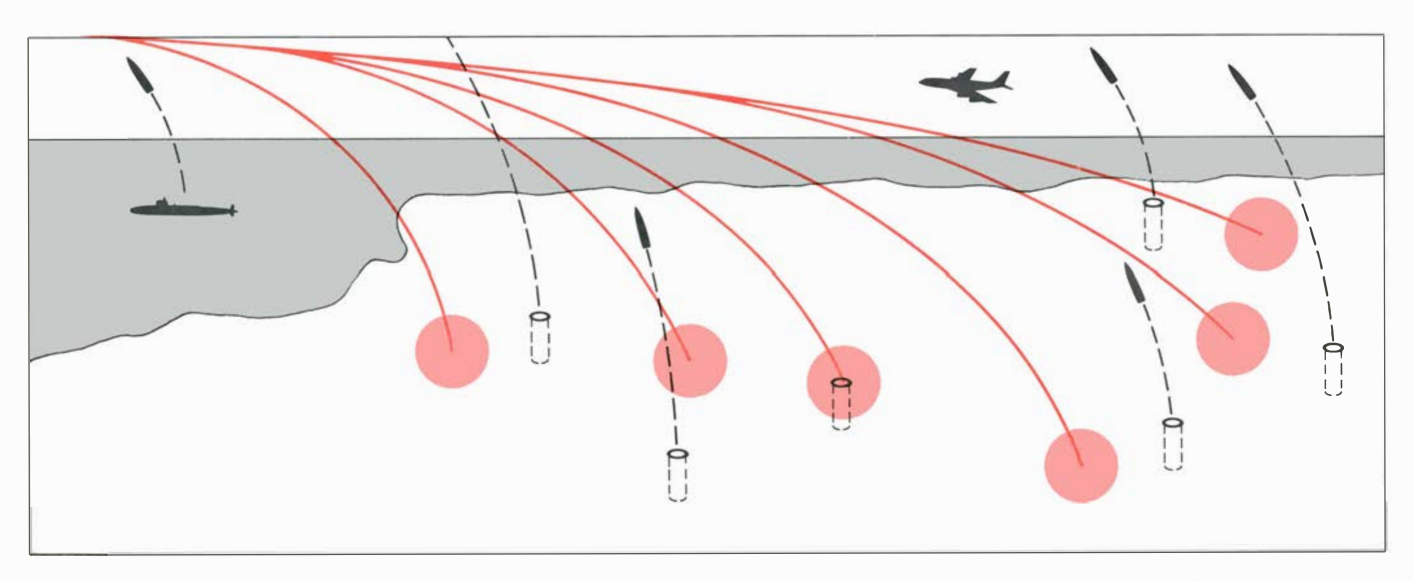




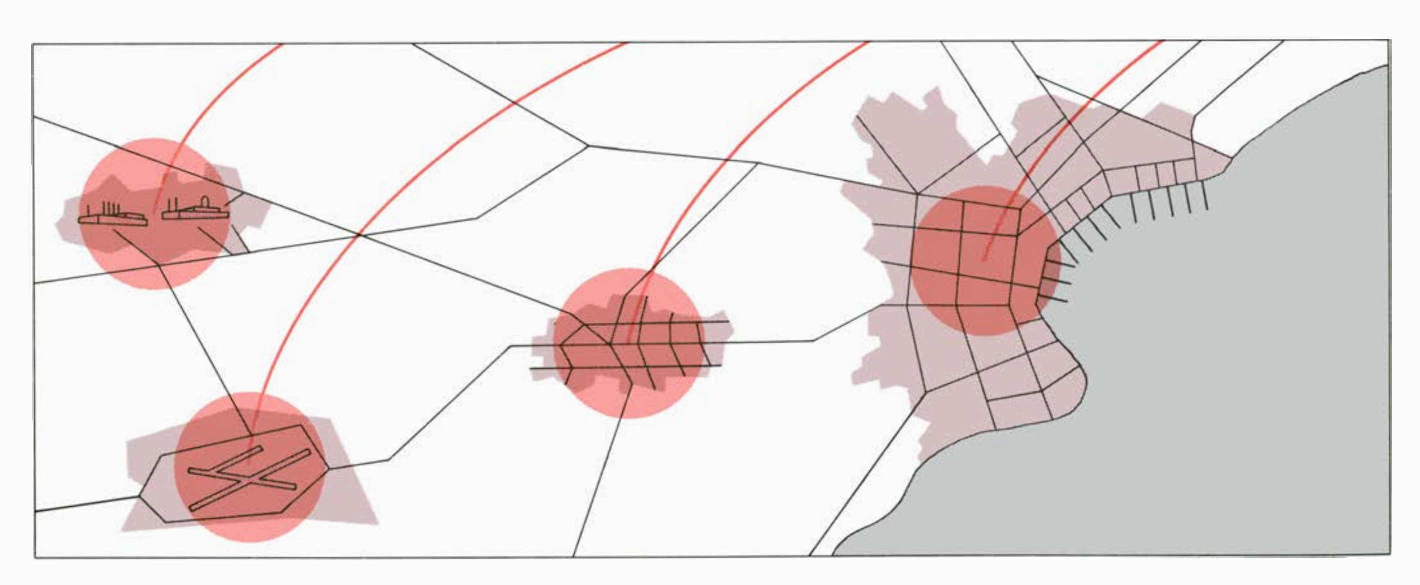
DE SU PRECISION Y DE SU POTENCIA EXPLOSIVA depende la probabilidad que tiene un misil de destruir un objetivo protegido. Una cabeza nuclear de poca potencia, pero con pequeño error circular probable (ECP), puede destruir un silo subterráneo (*a*), del mismo modo que otra con mayor potencia pero mayor ECP, o sea, menor precisión (*b*); cualquiera de estas dos cabezas nucleares constituye un arma de "contrafuerza" eficaz. Un arma con potencia o precisión insuficientes para poder destruir un silo no es ningún arma de "contrafuerza" eficaz (*c*).



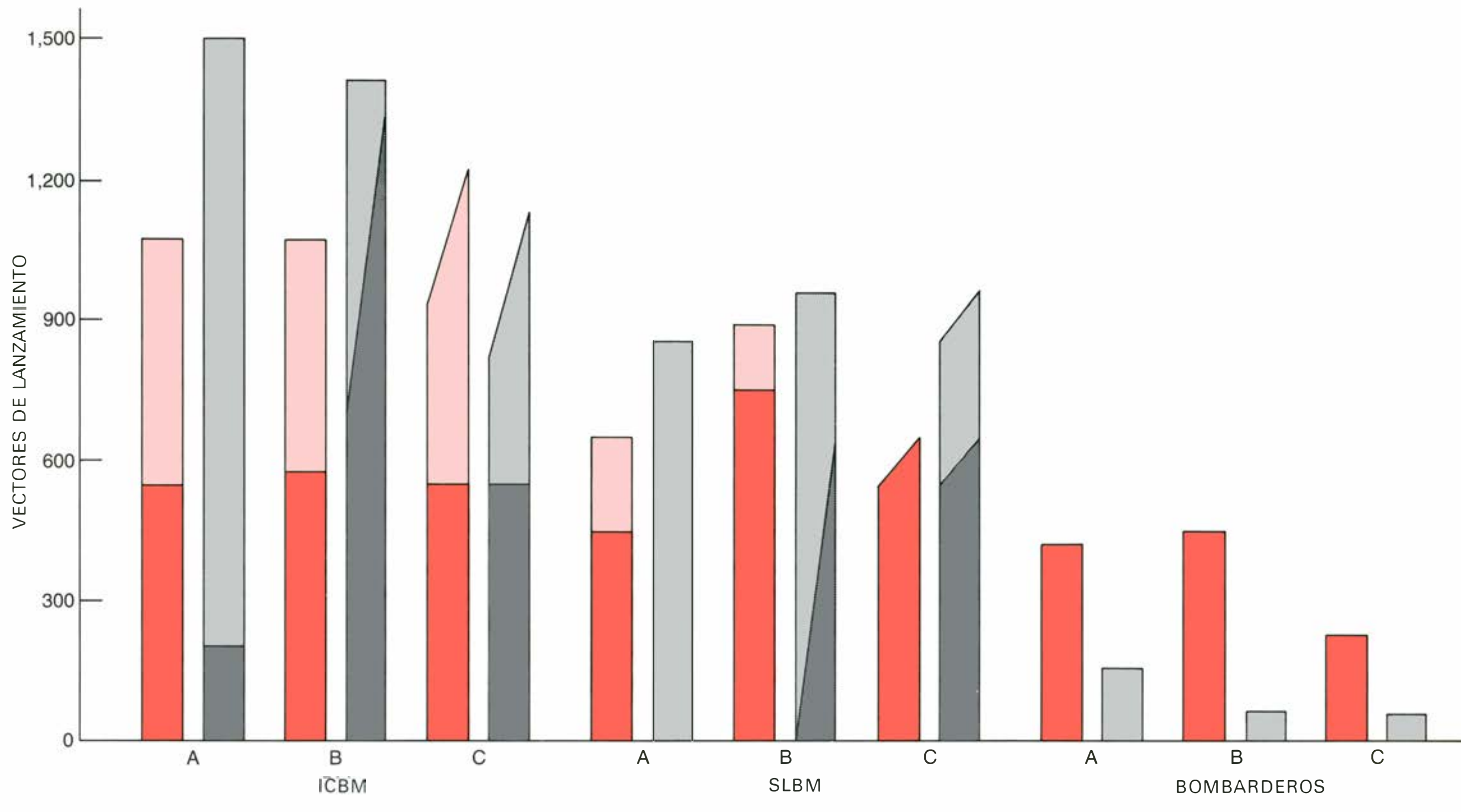
UN ATAQUE EFICAZ de "contrafuerza" con misiles dotados con ojivas de múltiples cabezas ("multiple independently targetable reentry vehicles": MIRV) destruye la mayoría de los misiles balísticos intercontinentales ("intercontinental ballistic missiles": ICBM). Bombarderos en estado de alerta y submarinos lanzamisiles permanecen todavía en condiciones de efectuar un ataque de represalia con plena eficacia.



UN ATAQUE FALLIDO de "contrafuerza" es aquel en el que los misiles atacantes carecen de la precisión o de la potencia explosiva necesarias para destruir muchos de los ICBM. La mayoría de éstos pueden ser lanzados y llevar a cabo un ataque de represalia en gran escala.



REPRESALIA EFICAZ efectuada por la nación que ha sido objeto de un ataque por sorpresa contra la que lo desencadenó. Los ICBM, SLBM ("submarine-launched ballistic missiles") y bombarderos atacan núcleos de población, puertos, objetivos industriales y otros objetivos militares que no sean silos subterráneos. Para una represalia eficaz se necesita una precisión inferior a la requerida para un ataque de "contrafuerza".



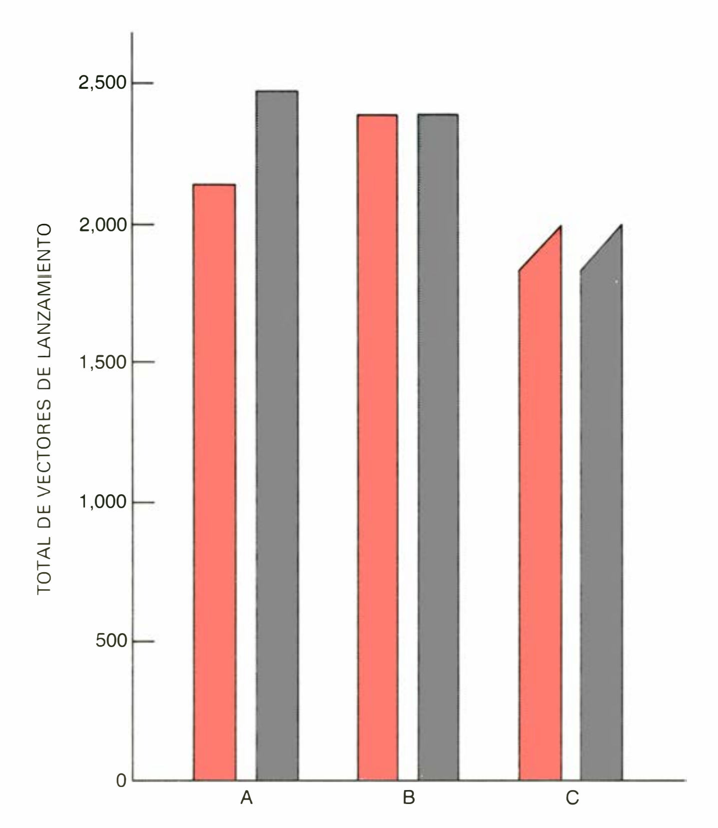
COMPARACION entre las fuerzas estratégicas de los Estados Unidos *(color)* y de la Unión Soviética *(gris)* de acuerdo con los datos vigentes en la primavera última *(A)*, según podrían ser en 1985 si fueran aceptados los techos de Vladivostok *(B)* o con los techos inferiores de la propuesta norteamericana comprehensiva *(C)*. Las cifras dadas ahora podrían variar ligeramente por cuanto proceden de hipótesis hechas sobre la forma en que las dos potencias pueden efectuar la distribución de sus medios, dentro del arsenal respectivo. El gráfico de la

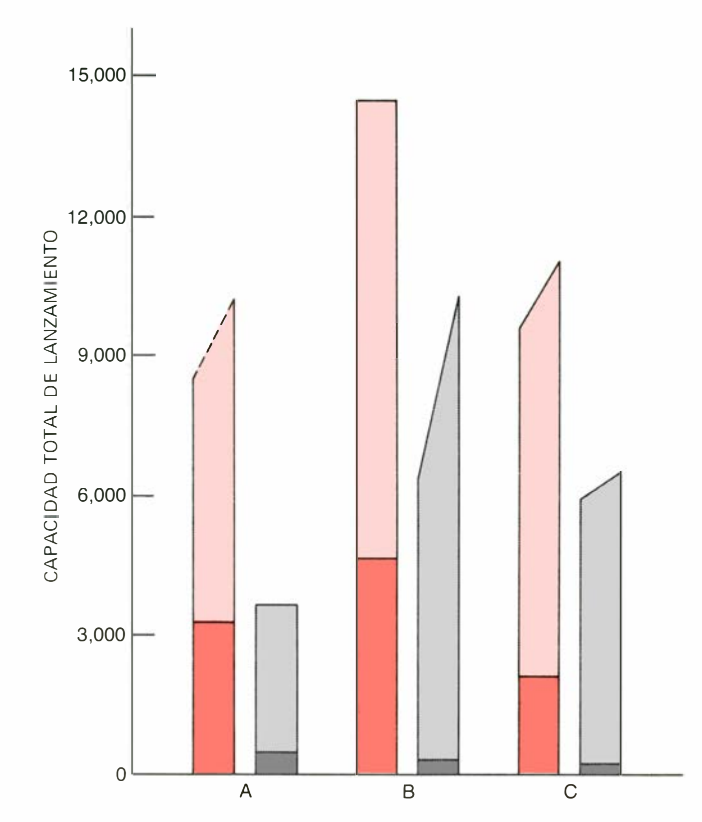
acuerdo provisional de 1972. La reducción en el número de ICBM sería presumiblemente más aconsejable que la de los costosos submarinos, que, además, constituyen la fuerza de disuasión que inspira más confianza. Si los soviéticos quisiesen –como así parece– alcanzar los techos permitidos en la propuesta, de 1100 a 1200 misiles con MIRV, el bajo límite señalado para cabezas nucleares múltiples en los ICBM haría que tuviesen que dotar con MIRV de 550 a 650 SLBM. Los soviéticos han comenzado a experimentar recientemente un SLBM con dos cabezas MIRV para el más moderno de sus dos tipos de submarinos nucleares lanzamisiles. Ese SLBM será muy inferior al Poseidón estadounidense, con MIRV de 10 cabezas, que está en servicio desde hace cinco años, y todavía más alejado del misil Trident I, previsto para hacerlo dentro de dos años. Por varias razones, cabe pensar, en consecuencia, que los rusos prefieran armar con MIRV un mayor número de ICBM. La limitación en cuanto al equipamiento con MIRV de los ICBM reflejada en la propuesta estadounidense fue considerada –comprensiblemente– por los soviéticos como parcial, en el sentido de que restringía las armas más valoradas por la Unión Soviética, y menos por los Estados Unidos.

Con 135 aviones que datan de la mitad de los años cincuenta, Rusia marcha muy por detrás de los Estados Unidos en lo que respecta a bombarderos intercontinentales. Sin embargo, no parece que esté haciendo muchos esfuerzos por acortar la ventaja. Si fuese necesario, cabe imaginar que los rusos podrían reducir el total general de armamentos estratégicos, con sólo retirar algunos de esos viejos bombarderos. Los Estados Unidos, por otra parte, posiblemente mantendrían un considerable número de B-52, cuya gran capacidad de carga sería un elemento significativo en el potencial nuclear que podría ser lanzado contra la Unión Soviética.

El nuevo bombardero soviético Backfire, que ha merecido una gran atención, tiene muy pocas posibilidades en misiones intercontinentales. Solamente puede alcanzar objetivos estadounidenses partiendo de bases rusas en el Artico, en vuelo subsónico a gran altitud y sin posibilidad de regreso. Una misión con retorno requeriría reabastecimiento en vuelo, técnica a la que los rusos nunca asignaron una prioridad digna de consideración. La capacidad del Backfire como bombardero estratégico –bombardero capaz de alcanzar el territorio enemigo– es ciertamente menos significativa que la de los bombarderos estadounidenses con base en Europa o en portaaviones, los cuales pueden alcanzar objetivos estratégicos en el interior de Rusia. Los Estados Unidos afirman que la misión primordial de sus aviones en "bases adelantadas" no es estratégica, sino de apoyo a sus aliados de la OTAN en caso de conflicto en Europa. Similarmente, los soviéticos mantienen que la utilización del Backfire está prevista contra objetivos periféricos (China o Europa occidental).

Existen, por tanto, varios caminos por los que podría llegarse a la reducción de las fuerzas estratégicas de ambos países para alcanzar los techos fijados en la propuesta comprehensiva. Sin embargo, las diferentes opciones se ven constreñidas por la estructura de los arsenales actuales, el nivel tecnológico que puede alcanzar cada país y los condicionantes burocráticos en ambos. Alcanzar los techos totales propuestos –de 1800 a 2000 vectores estratégicos– sería más difícil para los soviéticos, ya que les supondría una disminución de 500 a 700 vectores con respecto a los niveles actuales, mientras que para los norteamericanos esta disminución se podría fijar entre 150 y 350 vectores. Los Estados Unidos podrían proceder fácilmente a esta reducción retirando bombarderos o ICBM viejos, mientras que la Unión Soviética tendría





izquierda muestra las cifras de ICBM, SLBM y bombarderos estratégicos, con indicación del número de misiles que podrían ser dotados con MIRV *(color oscuro y gris oscuro).* **Las líneas inclinadas muestran los márgenes de posibilidades. El gráfico del centro muestra la cifra total de vectores estratégicos. La ilustración de la derecha, por su parte, representa la capacidad total de lanzamiento, que es la suma de la capacidad de lanzamiento con bombarderos estratégicos** *(tonos oscuros)* **y de la capacidad de lanzamiento con ojivas de los misiles** *(tonos claros).*

que retirar ICBM más modernos o, incluso, SLBM. El techo de 1100 a 1200 en cuanto al total de vectores con MIRV no supondría reducción alguna en los arsenales actuales de ambos países, pero la conclusión de que sólo 550 ICBM puedan llevar ojiva múltiple resultaría difícil de cumplir por parte de la Unión Soviética. Si los Estados Unidos se decidiesen a seguir adelante con proyectos como el del submarino Trident, la adopción de los techos más bajos les obligaría a retirar de servicio algunos de los misiles con MIRV ahora desplegados.

Caso de ponerse en práctica la propuesta comprehensiva, los rusos podrían alcanzar en 1985 una superioridad de dos a uno, con respecto a los norteamericanos, en cabezas nucleares con ICBM: la potencia de sus cabezas nucleares sería aproximadamente de un megatón, frente a los 200 kilotones que tendrían los estadounidenses. La ventaja soviética en el número de ICBM sería algo más pequeña que la derivada de los acuerdos de Vladivostok, pero sustancialmente más importante que la actual, ya que Rusia está empezando a poner en servicio las ojivas MIRV. La ventaja soviética en número y potencia de las cabezas nucleares se vería compensada por la superioridad norteamericana en precisión y, probablemente, en seguridad de empleo.

Por lo que respecta a cabezas montadas en SLBM, la superioridad norteamericana sería aproximadamente de dos o tres a uno. Además, los misiles y submarinos de los Estados Unidos son tecnológicamente superiores y cuentan con la ventaja geográfica de que sus navíos pueden operar con toda facilidad en el océano Atlántico. En materia de bombarderos, los Estados Unidos tendrían una supremacía aplastante, incluso en el caso de que el bombardero soviético Backfire fuese considerado como arma estratégica. La superioridad norteamericana en bombarderos y SLBM compensaría con creces la de los rusos en ICBM. Los Estados Unidos alcanzarían, en consecuencia, cierta ventaja en el número de cabezas nucleares, y otra ventaja muy superior en capacidad total de lanzamiento, entendiendo por tal la suma de cabezas nucleares en ICBM y SLBM y bombas nucleares de los aviones. Dicha ventaja, sin embargo, podría no ser tan grande como la potencialmente alcanzable si se aplicaran los acuerdos de Vladivostok. De cualquier forma, dicha superioridad no supone nada, militarmente hablando, ni tendría significación, ya que, en caso de un conflicto nuclear, ambos países cuentan con una potencia más que suficiente para destruir a su respectivo oponente.

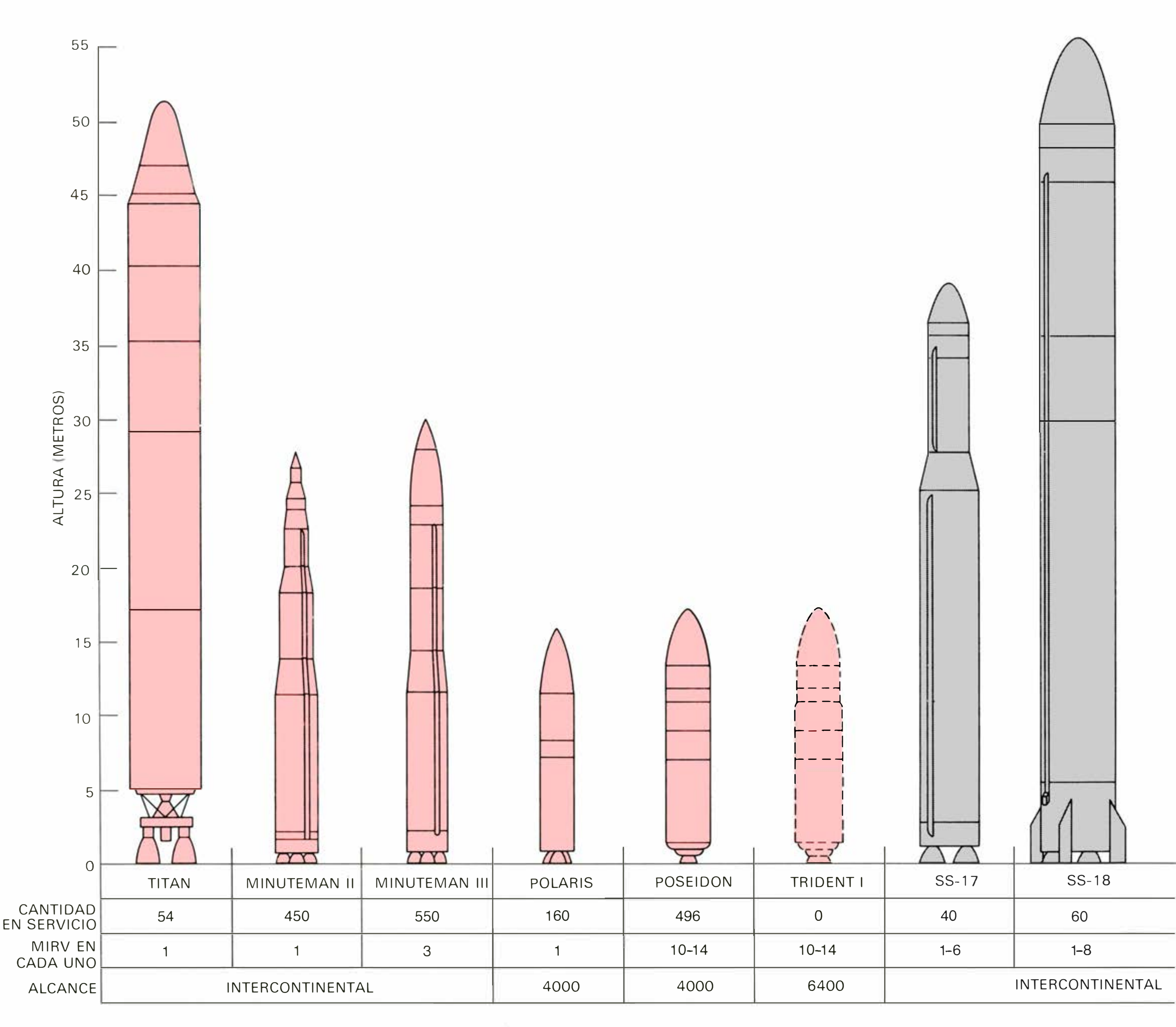
Lo que realmente cuenta es que si se consiguiesen los techos de la propuesta comprehensiva, ambas potencias quedarían, a grandes rasgos, equilibradas. Cada una detentaría la supremacía en los sistemas de armas en los que ha venido centrando sus esfuerzos en el pasado. Si fuesen congelados los avances tecnológicos (un condicional importante que depende de los techos cualitativos a los que me referiré más adelante), ninguno de los dos países lograría una capacidad de "contrafuerza" –que es la capacidad de destruir la mayoría de los misiles del otro bando– aceptable. Ambos dispondrían de una fuerza de disuasión segura y el equilibrio estratégico permanecería estable, ya que existirían escasos incentivos para desencadenar un ataque nuclear por sorpresa (nuclear first strike).

Las propuestas del presidente Carter sobre avances cualitativos en armamentos fueron potencialmente más trascendentales que las limitaciones en cuanto al número de armas. Desde hace algunos años, ha quedado patente que la carrera de armamentos estratégicos se ve impulsada, no por la necesidad de aumentar el número de armas, sino por la continua búsqueda de avances tecnológicos con vistas a lograr una superioridad. Ni el acuerdo provisional de 1972 ni los de Vladivostok establecieron medida algu-

na en este sentido: la sustitución de misiles por otros más modernos fue contemplada tácitamente o, en algunos casos, permitida explícitamente.

La propuesta comprehensiva prohibiría la modernización o sustitución de ICBM, así como el desarrollo y despliegue de ICBM móviles. Limitaría las pruebas en vuelo de misiles balísticos a seis por año para ICBM y otros seis para SLBM. Prohibiría los misiles crucero estratégicos –aquellos que tienen un alcance superior a los 2500 kilómetros (1550 millas)–, mientras que no habría restricción alguna para los misiles crucero de menor alcance. Finalmente, requeriría garantías por parte de los soviéticos en cuanto al Backfire, en el sentido de no darle capacidad intercontinental ni ser desplegado como medio estratégico. Las prohibiciones propuestas sobre modernización e ICBM móviles y las limitaciones en el número de pruebas experimentales tendrían por objeto, fundamentalmente, prevenir, o al menos retrasar por largo tiempo, cambios sustanciales en los arsenales actuales. En particular, las estrictas limitaciones sobre ICBM evitarían que ninguno de los dos países adquiriese una capacidad de "contrafuerza" que pudiese poner en peligro los ICBM del otro bando. Estas cláusulas incrementarían notablemente la estabilidad del equilibrio estratégico y podrían frenar de modo inmediato la carrera de armamentos.

La cuota de seis pruebas por año reforzaría, de forma eficaz, las prohibiciones sobre modernización y sustitución de misiles, ya que se necesitan muchas más pruebas para alcanzar una fiabilidad que asegure que las armas desplegadas son realmente operativas. A falta de pruebas adicionales, sería casi imposible introducir mejoras importantes en los misiles ahora existentes: la adquisición de armamentos totalmente nuevos estaría fuera de lugar. Utilizando medios de verificación, denominados oficialmente "medios técnicos nacionales de verificación", podría asegurarse que las cuotas fueran respetadas. Desde la primera prueba soviética de ICBM en 1957, los Estados Unidos han podido observar los ensayos con un alto grado de veracidad, desde fuera de la Unión Soviética. La combinación de reconocimientos desde satélites y observaciones electrónicas o



	TITAN	MINUTEMAN II	MINUTEMAN III	POLARIS	POSEIDON	TRIDENT I	SS-17	SS-18
CANTIDAD EN SERVICIO	54	450	550	160	496	0	40	60
MIRV EN CADA UNO	1	1	3	1	10-14	10-14	1-6	1-8
ALCANCE	INTERCONTINENTAL			4000	4000	6400	INTERCONTINENTAL	

MISILES BALISTICOS INTERCONTINENTALES de los Estados Unidos *(color)* y de la Unión Soviética *(gris)*. En cada misil figuran la cifra conocida o estimada de los desplegados, el número de cabezas nucleares múltiples (MIRV) y el alcance. Se incluyen tres ICBM y tres SLBM de cada potencia. Los Estados Unidos han anunciado que pronto pondrán en servicio, en los Minuteman III, la

con rayos infrarrojos desde la periferia de la Unión Soviética han proporcionado buen caudal de información sobre los lanzamientos y sobre las operaciones de reentrada. La calidad de dicho control podría mejorar y alcanzar mayor seguridad mediante un acuerdo por el que todos los lanzamientos fuesen realizados con arreglo a condiciones previamente determinadas, de igual modo a como se hizo con ocasión de negociarse el tratado de misiles antimisiles de 1972.

Estas propuestas de tipo cualitativo congelarían eficazmente los niveles tecnológicos que las dos naciones hubiesen alcanzado en el momento de la entrada en vigor de las limitaciones. Dos programas estadounidenses de ICBM en desarrollo se verían seriamente afectados.

SS-19	SS-N-6	SS-N-8	SS-NX-18
140	554	296	0
1-6	1	1	2–3
	2100 2600	6700	6700 (?)

nueva cabeza Mk, 12A, que tiene mayor potencia explosiva y un sistema de guiado más preciso. Trident I entrará en servicio en breve.

Los Estados Unidos están completando el desarrollo de la ojiva Mk. 12A, con sistema de guiado NS-20, que sustituirá a los MIRV que monta el misil Minuteman III; la entrada en servicio de este sistema de guiado está prevista para finales del presente año. La nueva ojiva mejorará la precisión del MIRV del Minuteman III y duplicará aproximadamente la potencia explosiva de sus cabezas nucleares, que tendrán una elevada probabilidad (más del 70 por ciento) de poder destruir un silo subterráneo. El misil Minuteman lograría así lo que los soviéticos podrían considerar como una seria capacidad de "contrafuerza". Los Estados Unidos han comenzado a diseñar también un ICBM totalmente nuevo, el misil MX, para reemplazar al Minuteman. El MX tendrá cabezas nucleares de mayor potencia con sistema de guiado terminal (esto es, dirigidas en la última parte de su trayectoria) hacia el objetivo. Se trataría de una verdadera arma de "contrafuerza" y, por tanto, desestabilizaría realmente el equilibrio. El MX va a ser móvil y, en consecuencia, muy difícil de destruir; podrá desplazarse en ambos sentidos a lo largo de una trinchera protegida o salir disparado desde cualquier rampa de lanzamiento preparada al efecto. Esto aumentará su capacidad de supervivencia y, al mismo tiempo, hará imposible que los rusos verifiquen la cantidad de misiles desplegados. Si los soviéticos hicieran otro tanto nunca más sería posible negociar ningún techo o límite sobre ICBM. Los dos programas citados quedarían interrumpidos de aceptarse la propuesta comprehensiva norteamericana.

Rusia se vería constreñida, de igual modo, por esta propuesta. Los nuevos ICBM rusos (los SS-17, SS-18, SS-19 y, muy probablemente, el SS-16) habrán sido ensayados lo suficiente como para poder ser desplegados con sus características actuales, pero no se cree que versiones de ellos con MIRV posean en estos momentos la precisión y potencia necesarias que les proporcione una verdadera capacidad de "contrafuerza". La propuesta norteamericana impediría su perfeccionamiento, lo cual redundaría en que los soviéticos no llegasen a disponer de una "contrafuerza" capaz de poner en peligro el arsenal de misiles Minuteman III. La prohibición de misiles móviles impediría que los soviéticos desarrollasen un modelo móvil del SS-16 o diesen capacidad intercontinental al misil móvil de alcance intermedio SS-20, dos posibilidades que han causado honda preocupación en Washington. En resumen, los Estados Unidos quedarían con vectores tecnológicamente más avanzados, lo cual estaría compensado con una mayor capacidad de carga de los misiles soviéticos actuales. En consecuencia, existiría cierto equilibrio en la capacidad estratégica total de que disfrutan ambas potencias.

En lo que respecta a SLBM, la situación sería diferente. Como ya he indicado, los Estados Unidos están en la actualidad muy por encima de Rusia en este campo. A menos que se incluyesen en la propuesta algunas cláusulas permitiendo a los soviéticos ponerse a su altura –por ejemplo, que completaran el desarrollo de cabezas MIRV de largo alcance para sus SLBM– no es probable que fuesen aceptadas las cuotas de seis pruebas por año. Si la limitación de pruebas entrase en vigor a corto plazo, la ventaja norteamericana se vería compensada de algún modo por el hecho de que la cuota haría imposible el despliegue del SLBM Trident I y el desarrollo del Trident II. La cancelación del Trident II detendría, por vía indirecta, el programa de submarinos Trident, ya que el desarrollo de tales submarinos no sería justificable si se limitase su armamento a los misiles Trident I, que pueden ser lanzados desde submarinos más pequeños. (Cancelar el programa de submarinos Trident no supondría una merma en cuanto a seguridad porque, en cualquier caso, los submarinos menores seguirían siendo más aconsejables.) Al ser los SLBM, básicamente, armas de disuasión más que de "contrafuerza", un acuerdo que demorase la entrada en vigor de la limitación de pruebas con SLBM hasta tanto no se completase el Trident I y los soviéticos dispusieran de SLBM con MIRV, no supondría un daño considerable para el control de armamentos. Podría, incluso, incrementar la estabilidad del equilibrio de disuasión.

Parte importante de la propuesta comprehensiva es la que trata del misil crucero que ha estado en desarrollo en los Estados Unidos –pero no en la Unión Soviética– desde 1972. Se trata de una pequeña y económica "bomba volante" de largo alcance, derivada de la V-1 de la Segunda Guerra Mundial, muy mejorada gracias a componentes electrónicos miniaturizados, mediante los cuales puede batir objetivos concretos con extraordinaria precisión (véase "Los misiles crucero", por Kosta Tsipis, en INVESTIGACION Y CIENCIA, abril 1977). Los rusos han insistido en que los acuerdos de Vladivostok incluían, dentro de los techos de vectores estratégicos, a los misiles crucero con un alcance superior a 600 kilómetros (375 millas). Sin embargo, la propuesta comprehensiva estadounidense

define como estratégicos únicamente a los misiles crucero con un alcance superior a los 2500 kilómetros (1550 millas). Tal distancia es más que suficiente para alcanzar objetivos estratégicos de la Unión Soviética al oeste de los Urales, desde bases terrestres en Europa, submarinos u otros buques de superficie, o desde aviones en vuelo sin necesidad de penetrar en territorio soviético. No debe extrañar, por tanto, que los soviéticos pusieran en duda la sinceridad de la propuesta de los Estados Unidos basada en esta definición, teniendo en cuenta que los propios norteamericanos han estado llamando misiles crucero estratégicos a varias armas cuyo alcance era inferior a 2500 kilómetros.

Portavoces de la Administración norteamericana han argumentado que los misiles crucero con un alcance de hasta 2500 kilómetros no se diferenciaban de otros "sistemas de bases adelantadas" (los aviones de los Estados Unidos estacionados en Europa) por lo que, de igual modo, tampoco deberían ser prohibidos por un acuerdo SALT. El problema de tal argumentación es que vuelve a plantear la cuestión de la no inclusión en el acuerdo de los "sistemas de bases adelantadas", concesión importante que hicieron los soviéticos en el acuerdo provisional de las SALT I y también en Vladivostok. La ventaja norteamericana al no incluir este tipo de aviones dentro de los vectores estratégicos es patente, ya que mientras ellos disponen de 500 aviones de este tipo, la Unión Soviética no tiene ninguno (a menos que el Backfire sea considerado como tal). En compensación a la aquiescencia soviética a continuar excluyendo de las negociaciones SALT los "sistemas de bases adelantadas", lo mejor sería excluir el Backfire de las SALT (a menos que se le conceda verdadera capacidad intercontinental) e incluir en ellas aquellos misiles crucero cuyo alcance no es lo suficientemente corto como para que puedan ser considerados verdaderamente tácticos. El Backfire, los misiles crucero tácticos y la aviación instalada en "bases adelantadas" podrían ser abordados en el contexto de las negociaciones MBFR ("mutual balanced force reduction": reducción mutua y equilibrada de fuerzas) que han estado desarrollándose en Viena, durante los tres últimos años, entre los países del Pacto de Varsovia y la OTAN.

La propuesta comprehensiva, rechazada de plano por los soviéticos en Moscú, era obviamente demasiado extensa, e incluía demasiados conceptos nuevos y delicados como para que pudiera servir de base a un acuerdo en las SALT antes de la fecha de expiración en este mes de octubre. Ya que los Estados Unidos presentaron como alternativa la propuesta "expectante", era lógico concluir combinando las dos para llegar al acuerdo de trabajo sobre tres puntos descritos anteriormente: un tratado por cinco años con arreglo a los niveles o techos de Vladivostok o un poco más rebajados; un protocolo de tres años de validez, restringiendo de algún modo los misiles crucero, los ICBM pesados armados con MIRV y el Backfire; por último, una declaración de principios para llegar a reducciones eficaces y a limitaciones cualitativas en el futuro. La cuestión más importante y el aspecto más delicado, entre los dos países, puede que sea la manera concreta de restringir el misil crucero.

Un posible enfoque con vistas a tal acuerdo está contenido en una propuesta que hicieron los rusos durante las conversaciones de enero de 1976, en Moscú. Propusieron la prohibición de los misiles crucero lanzados desde mar o desde tierra con un alcance superior a 600 kilómetros, mientras que los misiles crucero lanzados desde el aire con un alcance de hasta 2500 kilómetros no quedarían sujetos a limitación alguna, siempre que los aviones con capacidad de lanzar esos misiles quedasen incluidos dentro de los 1350 vectores estratégicos con MIRV, acordados en Vladivostok. Como en la actualidad hay 1046 misiles de los tipos Minuteman III y Poseidón con MIRV y cada uno de los diez o más submarinos Trident programados llevará 24 misiles MIRV, los Estados Unidos dispondrán de un total de 1286 vectores con MIRV. Esto significaría que únicamente 34 aviones estadounidenses podrían ir armados con misiles crucero. No obstante, éste podría ser el punto de partida para llegar a un acuerdo sobre misiles crucero. La idea de contabilizar como vectores MIRV a los aviones equipados para lanzar misiles crucero es acertada, ya que impide una ilimitada entrada en servicio de misiles crucero lanzados desde el aire. También, porque coloca la cifra de vectores MIRV –bastante elevada– dentro de unas medidas más significativas a efectos de control de armamentos y permite, sin embargo, que algunos aviones se utilicen como vectores estratégicos sin tener que penetrar en las defensas aéreas soviéticas.

De todas formas, habría problemas en la verificación. El alcance operativo de un misil crucero no puede ser constatado con certeza mediante la observación de las pruebas que realice el contrario; también puede resultar difícil verificar que un misil de largo alcance, declarado para ser lanzado desde el aire, no lo sea desde tierra o desde un navío, lo cual violaría los inferiores límites de alcance señalados para dichas armas. Sin embargo, los soviéticos nunca se han preocupado gran cosa por las verificaciones, ya que es virtualmente imposible para los Estados Unidos guardar en secreto cualquiera de sus grandes programas militares. Desde el punto de vista estadounidense, el hecho de que la Unión Soviética comenzase a desarrollar modernos misiles crucero de mayor alcance –aunque dicho alcance supere los 600 kilómetros– no llegaría a suponer una seria amenaza estratégica. Ni siquiera un misil crucero terrestre con un alcance de 2500 kilómetros podría penetrar en los Estados Unidos si fuese lanzado desde territorio ruso. Colocar dichos misiles en submarinos no tendría mucho valor militar para la Unión Soviética ni supondría un riesgo para la seguridad de los Estados Unidos, ya que los soviéticos poseen en la actualidad 840 misiles balísticos de mayor alcance contra los que Estados Unidos carece de defensa. Por otra parte, esto implicaría un uso poco adecuado de la fuerza submarina soviética y no merecería la pena violar un tratado actuando así. Si los soviéticos instalasen misiles crucero en los Backfire para darles capacidad intercontinental, esos bombarderos se contabilizarían dentro del techo de vectores con MIRV.

Los soviéticos han indicado que no aceptarán las considerables reducciones con respecto a los techos de Vladivostok, que supone la propuesta norteamericana comprehensiva, pero sí podrían acceder a reducciones inferiores, tal vez de un 10 por ciento en lugar del 25 por ciento. Esto no tendría mayor importancia en cuanto al control de armamentos, pero sería muy significativo para dar a entender al resto del mundo que las dos superpotencias nucleares estaban, por vez primera, dispuestas a frenar la carrera de armamentos. Por lo que respecta al control cualitativo, sería probablemente demasiado optimista pensar que en este mes de octubre se establezcan limitaciones en cuanto a modernización o sustitución de los armamentos hoy existentes. Ahora bien, acuerdos de este género podrían tomar forma a título de metas, en una declaración de principios para futuras negociaciones. Estas deberían proseguir sin demora una vez sea firmado el tratado SALT II. Cabe pensar en que ambos países se declaren dispuestos a efectuar, independientemente, restricciones en sus programas de armamentos durante futuras negociaciones, estableciendo así un clima favorable para empezar a poner bajo control la carrera cualitativa de armamentos.

Objetos BL Lacertae

Se encuentran entre los objetos del universo más alejados y luminosos. La fuente de su radiación es un pequeño núcleo (del tamaño del sistema solar), cuya luminosidad pudiera superar incluso a la de una galaxia

Michael J. Disney y Philippe Véron

Los objetos más curiosos del firmamento brillan con una luz totalmente diferente de la que desprenden las estrellas y las agregaciones de estrellas conocidas. Varias galaxias elípticas gigantes radian prodigiosamente en longitudes de onda de radio en las que ni las estrellas ni las galaxias normales lo hacen. Las nebulosas espirales, llamadas galaxias Seyfert, algunas de las cuales son también emisoras de radio, tienen regiones del núcleo cuya radiación intensa y espectralmente característica no puede explicarse como luz estelar. Los más extraordinarios de todos son los "faros cósmicos" (cosmic beacons) llamados objetos cuasi-estelares o quasars. Los quasars pueden estar enormemente alejados, de forma que la luz que hoy nos llega desde algunos de ellos, según parece, ha estado en camino durante la mayor parte de la historia del universo. Estas grandes distancias implican luminosidades lejos de toda comparación con las observadas en nuestra propia galaxia o en galaxias próximas. En un quasar, la radiación, muchas veces más potente que la de una galaxia entera, parece emanar de una región nuclear compacta que puede ser, aproximadamente, del tamaño del sistema solar.

En los últimos años, otra clase de objetos ha venido a sumarse a este catálogo de prodigios astronómicos: los objetos BL Lacertae. No hace ni 10 años que se descubrió el primero de ellos y hoy en día ya se conocen más de 30. Al igual que los quasars, parecen enormemente alejados y son inexplicablemente brillantes, pero tienen peculiaridades espectrales comunes. Como en los quasars y galaxias Seyfert, su intensa radiación procede, según se deduce, de un volumen relativamente pequeño.

En un principio, el descubrimiento de los objetos BL Lacertae sirvió para enredar más la confusión reinante sobre la naturaleza de muchas nuevas fuentes luminosas, que por entonces parecían proliferar sin límites. Hoy ha comenzado a esclarecerse la posibilidad de que exista una continuidad de propiedades comunes a todos estos objetos extraños. No conocemos todavía la fuente última de su energía ni cómo la energía se convierte en radiofrecuencia y radiación visible. Pero podemos comenzar ya a clasificar la profusión y variedad de objetos en un esquema común. Cabe la posibilidad de que sean manifestaciones plurales de un proceso único operando en diferentes medios.

Todos los objetos BL Lacertae han sido reconocidos por sus emisiones en radiofrecuencias. El primero de ellos se puso de manifiesto cuando J. M. Macleod y B. H. Andrew, del Canadian National Radio Observatory, dirigieron un radiotelescopio de alta frecuencia a una conocida fuente de radio designada por VRO 42.22.01. Les había interesado esa fuente porque las observaciones en frecuencias de radio (varios cientos de megahertzios) habían mostrado que su espectro de radio presentaba una configuración insólita. En la mayoría de las fuentes de radio, la intensidad de la radiación disminuye al aumentar las frecuencias (o lo que es equivalente, a longitudes de onda más cortas). El espectro de radio de un quasar suele ser casi plano; la intensidad permanece casi constante en todas las longitudes de onda de radio. Se observó que, en VRO 42.22.01, la relación usual se invertía: la intensidad de la señal se incrementaba a las frecuencias más altas. Una simple extrapolación de esta tendencia sugería que a frecuencias superiores incluso, VRO 42.22.01 debía ser una fuente bastante intensa. Macleod y Andrew decidieron comprobar esta hipótesis en la frecuencia de su instrumento, que era de 10.600 megahertzios (equivalente a una longitud de onda de 2,8 centímetros). A esa frecuencia, VRO 42.22.01 volvía a ser uno de los 30 objetos que más brillaban en el cielo.

Hasta entonces no se había podido identificar a VRO 42.22.01 con ningún objeto discernible en longitudes de onda óptica, debido en parte a que se halla situado en la dirección de la Vía Láctea, lo que complica la búsqueda de su imagen óptica. En esa parte del cielo, las estrellas son abundantes en primer plano, y, junto con las densas nubes de polvo y gas, hacen confusa nuestra visión acerca de lo que se encuentra a distancias mayores en el universo. No obstante, Macleod y Andrew refinaron sus observaciones de radiofrecuencia para obtener una fijación precisa de la posición de la fuente. Cuando esa posición pudo comprobarse en fotografías del firmamento, se constató que coincidía con la imagen de una estrella de magnitud 14.

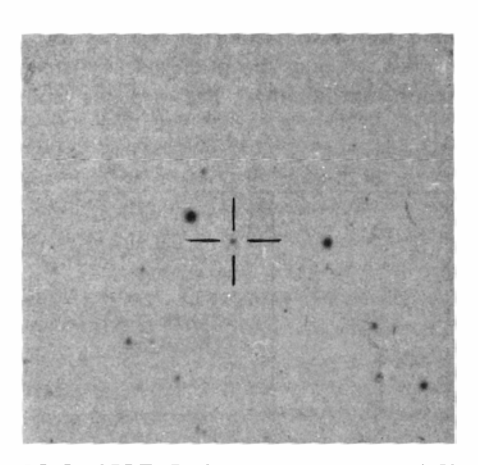
Una estrella de magnitud 14 es un objeto difícilmente visible con los medios de que dispone un observador casual; está tan alejada y es tan débil que no puede detectarse a simple vista. (Los incrementos de magnitud óptica corresponden a disminuciones de brillo. La escala es logarítmica y está calibrada de forma que un cambio de cinco magnitudes represente un cambio de cien veces en la luminosidad aparente). Macleod y Andrew se sorprendieron al dar con un objeto de este brillo. Todas las fuentes de radioemisión conocidas eran galaxias activas o quasars; entre esos objetos sólo un quasar brillaba más que la magnitud 14. Por otro lado, la localización casual de VRO 42.22.01 en la Vía Láctea permite que gran parte de su luz sea absorbida por el "humo" de nuestra propia galaxia, de lo que se desprende que su verdadera luminosidad es mucho mayor todavía.

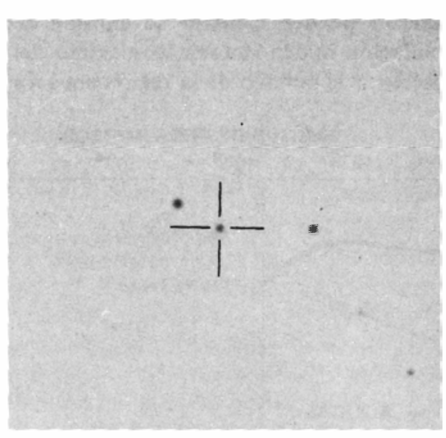
Diversos indicios sobre la naturaleza de VRO 42.22.01 estaban ya implícitos en las observaciones iniciales. En las longitudes de onda de radio, la fuente permanecía sin resolver (era, en otras palabras, puntual), de lo que se desprendía que resultaba muy pequeña, se hallaba demasiado lejos, o ambas cosas. Quizá la pista más importante fue que las emi-

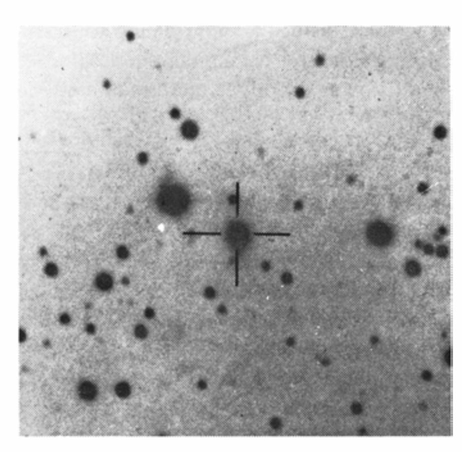
siones de radio estaban polarizadas. La polarización intrínseca de una fuente indica que la radiación no se genera por mecanismos térmicos, tales como la radiación desde una superficie caliente de una estrella. Además, la interacción de la radiación polarizada con el medio interestelar de nuestra galaxia, que gira el plano de polarización por diferentes cuantías a distintas frecuencias, nos facilita una medida grosera de la distancia: a partir de la rotación se puede estimar qué fracción de la galaxia ha cruzado la radiación. La observación de VRO 42.22.01 y su imagen óptica sugieren que la fuente se encuentra muy distante, probablemente fuera de nuestra galaxia. Por último, esa conclusión fue también corroborada por la aparición óptica del objeto. En las fotografías de larga exposición, una débil franja rodea al núcleo del objeto. Una hipótesis obvia es que la franja conste de estrellas y que VRO 42.22.01 sea una galaxia alejada. A partir de esta información y del espectro de radio invertido, Macleod y Andrew propusieron considerar, de forma provisional, que VRO 42.22.01 podría pertenecer a una clase particular de quasars y galaxias activos. Característica distintiva de esa clase es que sus emisiones son variables en las frecuencias ópticas y en las radiofrecuencias.

La hipótesis de trabajo propuesta por Macleod y Andrew fue seguida por J. L. Schmitt, del David Dunlop Observatory de la Universidad de Toronto, que emprendió una investigación de los registros de las estrellas variables conocidas. Descubrió que el objeto de Macleod y Andrew había sido ya identificado como una estrella variable en la constelación Lacerta (lagarto); se le había designado por BL Lacertae o BL Lac. La variabilidad había sido descubierta por el astrónomo alemán C. Hoffmeister en 1929, quien demostró que la estrella aumentaba de brillo y se apagaba a intervalos irregulares. Se observó un cambio del 100 por ciento en la luminosidad, en una semana; asimismo, se vio que, en el transcurso de un mes, el brillo podía cambiar en un factor de 15. En su brillo máximo BL Lacertae refulgía más que cualquier quasar.

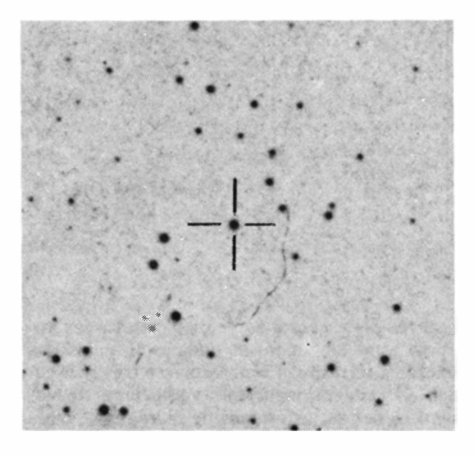
Los problemas de observación más interesantes, y los más difíciles de resolver, en astronomía suelen ser la determinación de la distancia, del tamaño y de la cantidad de energía que radia un objeto. Al no poder medir directamente estas cantidades, los problemas deben afrontarse a través de la única fuente de información útil: la débil radiación electromagnética que llega a la Tierra. La astronomía es el arte de extraer información de estas señales. La información estructural contenida en una imagen fotográfica es sólo el comienzo. También puede obtenerse información del espectro o distribución de energía radiante en función de la longitud de onda, del estado de polarización y de las variaciones en la luminosidad.

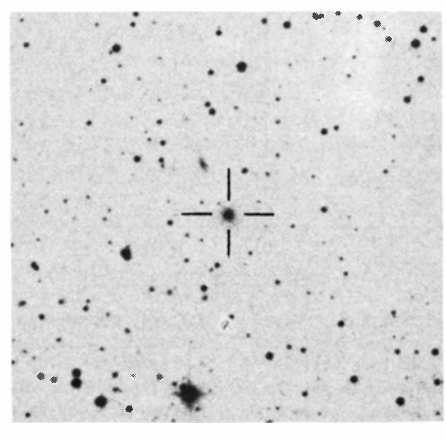


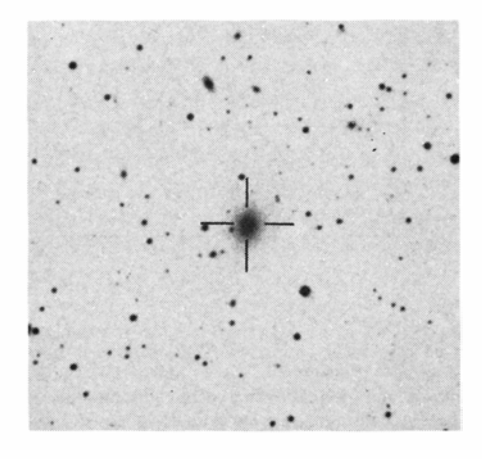




BL LACERTAE. Catalogado durante casi 50 años como una estrella variable, se le considera hoy una nueva clase de objetos astronómicos. La variabilidad es evidente en las dos fotografías de la izquierda, realizadas en el Wise Observatory de la Universidad de Tel-Aviv. La primera fotografía, empezando por la izquierda, se tomó el 25 de septiembre de 1973; en la del centro, sacada dos meses después, el 17 de noviembre, BL Lacertae (marcado por el retículo) es unas cuatro veces más brillante. En ambas fotos, BL Lacertae parece una estrella; a mayores exposiciones, se pone de manifiesto un halo borroso en torno a un núcleo brillante. En la foto de la derecha, tomada por G. Wlérick y D. Michet, puede distinguirse un halo un poco confuso. Su contorno sugiere la forma peculiar que posee una galaxia elíptica gigante.







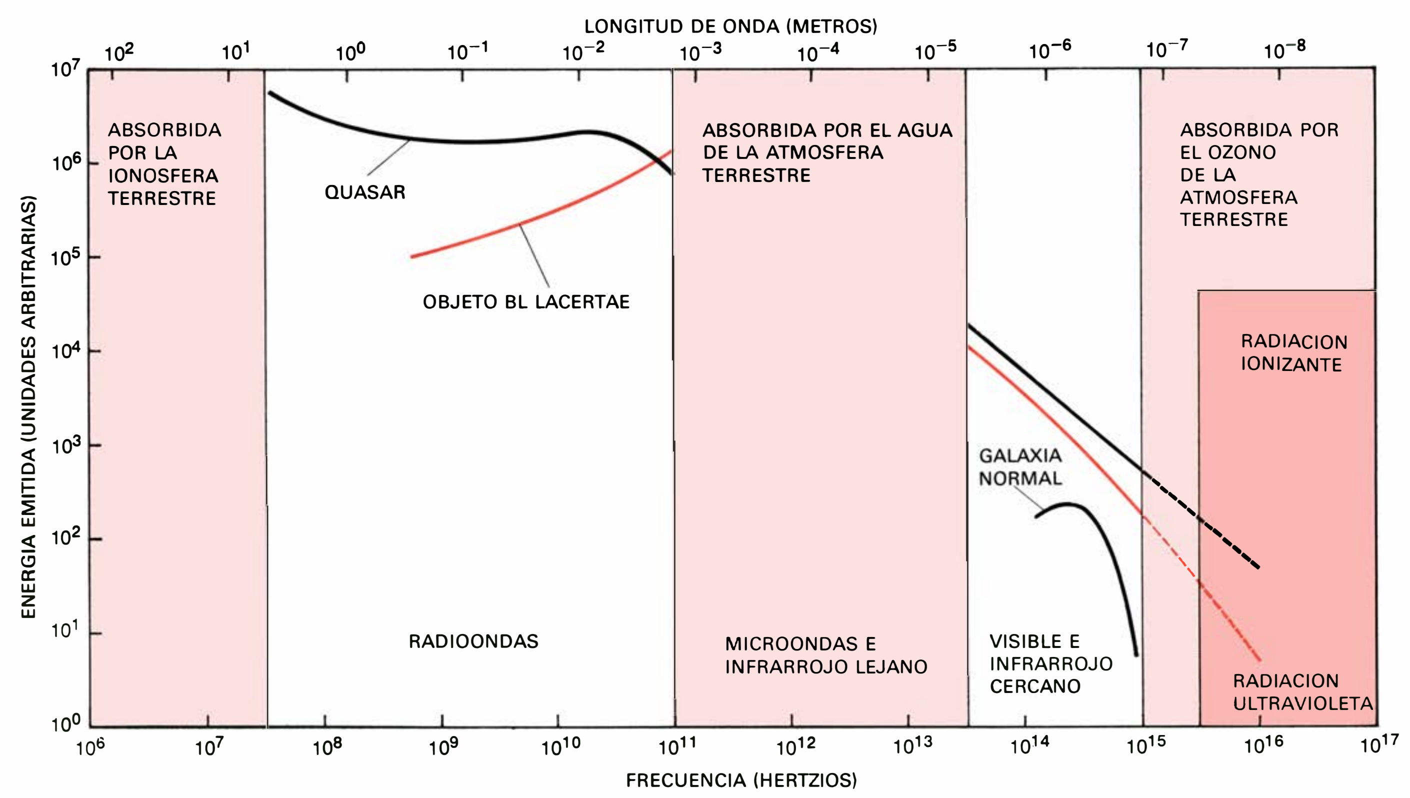
TRES OBJETOS, de distinta apariencia, que están todos ellos clasificados como objetos del tipo BL Lacertae. OJ287 *(izquierda)* forma una imagen completamente estelar, en donde no se distingue ninguna nebulosa de estrellas. AP Librae *(centro)* parece ambas cosas: un núcleo inequívoco y un halo galáctico; guarda una estrecha similaridad con BL Lacertae. Markarian 501 *(derecha)* tiene una nebulosa prominente y un núcleo tan débil que, sobre placas de larga exposición como ésta, la región del núcleo estrellado se halla oscurecida. Las tres fotografías se tomaron dentro del programa de trabajo del National Geographic Society-Palomar Observatory Sky Survey, de Estados Unidos de América.

El espectro óptico de BL Lacertae constituyó una gran sorpresa; en efecto, fueron principalmente las peculiaridades del espectro lo que llevó al establecimiento de una nueva categoría de objetos astronómicos. El espectro de una galaxia normal, cuya luminosidad es simplemente el brillo combinado de muchas estrellas, se distingue por una serie de líneas oscuras de absorción. Estas se producen cuando la radiación térmica de las estrellas, que tiene un espectro continuo, pasa a través de la atmósfera estelar de gas, que es frío en comparación. Los átomos de un gas pueden absorber sólo radiaciones de longitudes de onda específicas, en particular las longitudes de onda cuya energía corresponde a la diferencia energética entre estados cuánticos de los átomos. Las líneas de absorción generadas de un modo similar pueden detectarse, aunque débilmente, en el espectro de algunos quasars. Mucho más importantes en el espectro de los quasars son las líneas anchas de absorción, que destacan sobre el fondo, bastante más oscuro. Se supone que las líneas de emisión se producen cuando una nube de gas se calienta por una intensa fuente de radiación ultravioleta. Tales líneas constituyen una característica definitoria de los quasars. Las líneas de absorción y emisión en un espectro son también el medio que se emplea para estimar la distancia a que se encuentran galaxias y quasars.

Cuando se examinó el espectro óptico de BL Lacertae no pudieron registrarse ni líneas de absorción ni líneas de emisión. El espectro no presentaba rasgos distintivos, siendo por tanto inescrutable. En ausencia de líneas que marquen las conocidas transiciones atómicas de frecuencia, el espectro no ofrece indicio alguno sobre la distancia del objeto. En teoría, BL Lacertae puede no estar a distancia extragaláctica sino formar una estrella dentro de nuestra propia galaxia.

Otro resultado de interés se obtuvo al registrarse con un detector fotoeléctrico la variabilidad de BL Lacertae. La medida depende estrechamente del tamaño del objeto emisor. Los periodos más cortos en que una estrella o una galaxia pueden cambiar su emisión de radiación miden el tamaño máximo del objeto o el tamaño de la región emisora del mismo que está sujeta a variación. Ello se debe a que, en general, los cambios de brillo solamente pueden conseguirse por la propagación de cualquier señal a través de la región. La señal no puede viajar a mayor velocidad que la luz y, por tanto, la región variable no puede tener dimensiones superiores a la distancia que recorre la luz durante el periodo de variación más corto. Se vio que BL Lacertae variaba su brillo, como máximo en un 30 por ciento, en una sola noche. Esta relación implica que la mayor parte de la luz es emitida por una región cuyo diámetro no excede unos cuantos días luz. A escala, el tamaño del objeto se halla más próximo al de una estrella que al de una galaxia, la cual se extiende por varias decenas de miles de años luz.

Estas propiedades de BL Lacertae resultan paradójicas. Por una parte, las emisiones de radio y su polarización argumentan a favor de que el objeto sea extregaláctico, como una radiogalaxia o un quasar. Así lo confirma el halo de polvo visible en algunas fotografías, en que sólo pudieron ser detectados con dificultad. Por otra, las rápi-



EL ESPECTRO DE RADIACION revela grandes diferencias entre la energía total emitida por una galaxia normal y la emitida por objetos cuasi-estelares o la de un objeto BL Lacertae. Casi toda la radiación de una galaxia normal está en longitud de onda óptica, y el espectro tiene la configuración característica de la radiación térmica, como en la luz estelar. La radiación de los quasars y de los objetos BL Lacertae cubre una gama mucho más amplia de longitudes de onda. Además, la intensa emisión óptica de ambos radian fuertemente en las longitudes de onda de radio, siendo el espectro de los quasars casi plano en esta región, e, "invertido" e inclinado en una dirección opuesta a la normal, el de BL Lacertae. Tanto los quasars como los objetos BL Lacertae deben ser más potentes en las partes del espectro correspondientes a las microondas y al infrarrojo, pero estas longitudes de onda están eclipsadas por la absorción de la atmósfera terrestre. La absorción terrestre oscurece también las emisiones ultravioleta, incluyendo radiaciones de longitudes de onda corta, capaces de ionizar átomos de hidrógeno. Nubes de tal gas ionizado crean líneas de emisión en el espectro de los quasars, pero dichas líneas no aparecen en los espectros de BL Lacertae. La extrapolación desde la parte accesible del espectro sugiere que los Lacertae emiten una radiación ionizante relativamente pequeña; de forma alternativa, las líneas pueden estar ausentes, pues las galaxias elípticas asociadas a los objetos lacertae resultan deficientes en gas de hidrógeno.

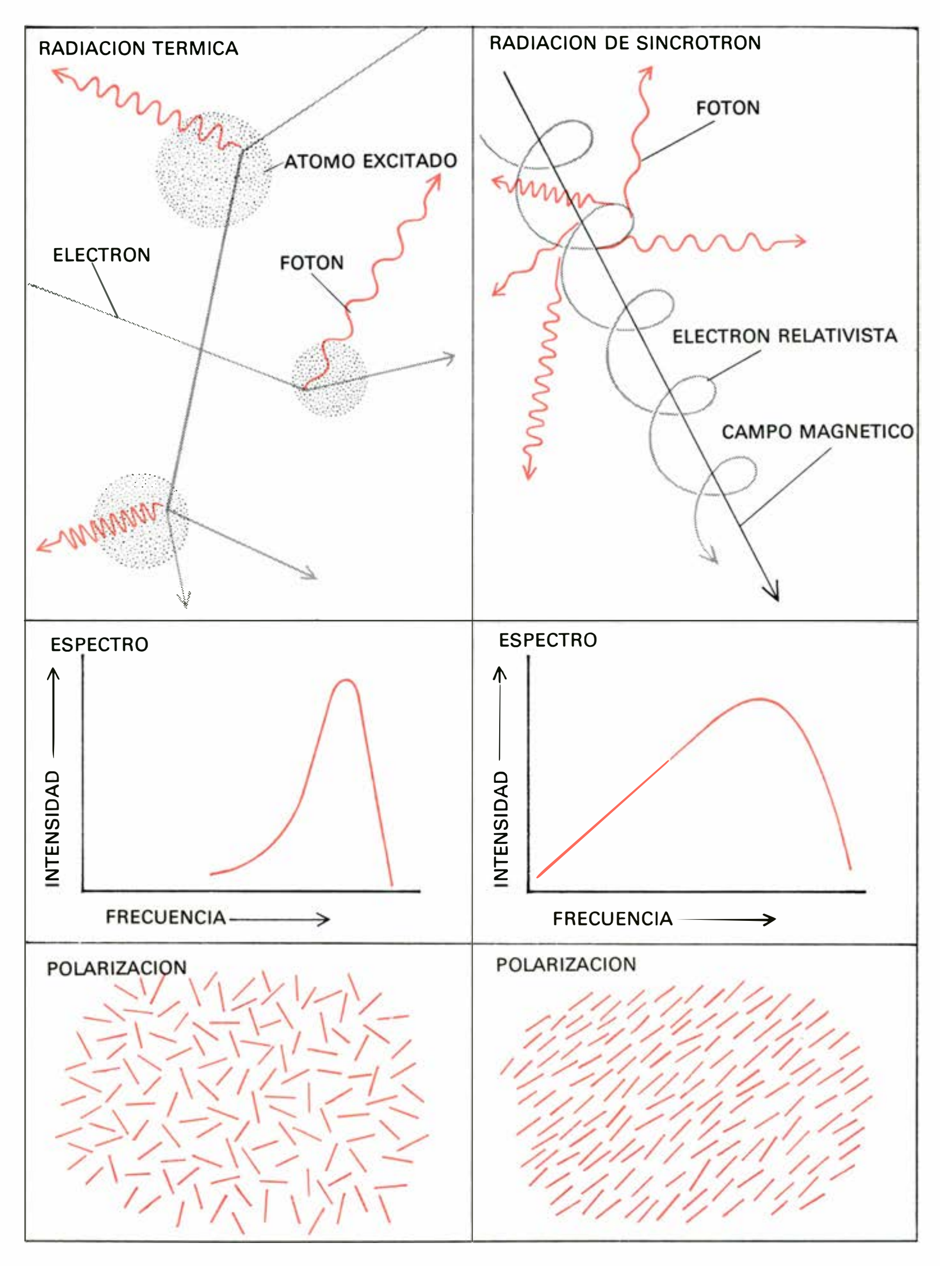
das variaciones en brillo sugieren una fuente pequeña, como pudiera serlo una estrella dentro de nuestra propia galaxia. La localización de la fuente en el interior de la Vía Láctea refuerza la posibilidad de la segunda hipótesis. Ninguna de las dos, sin embargo, podría explicar la falta de rasgos del espectro de BL Lacertae. La configuración del espectro –en otras palabras, el color del objeto– resultaba también difícil de explicar. Aun admitiendo cierta tolerancia, el objeto es demasiado rojo para poder asimilarlo a un quasar, dada la dispersión preferencial de la luz azul en el humo galáctico, pero tampoco es el espectro de una estrella.

No podía, pues, clasificarse a BL Lacertae dentro de las categorías existentes, y así se convirtió en arquetipo de una categoría nueva. Los aproximadamente 30 objetos similares que han sido identificados se encontraron más o menos de la misma forma: buscando fuentes puntuales brillantes de energía de alta radiofrecuencia. Habían pasado desapercibidos, ya que las primeras investigaciones en radioastronomía se hacían a frecuencia más baja. La denominación "objetos BL Lacertae" se ha venido aplicando a todos los miembros de la clase y aquí lo abreviaremos en "lacertae". Por supuesto, los objetos no suelen estar relacionados con la constelación Lacerta.

El segundo lacertae que se encontró se trataba de una radiofuente, designada por OJ287; como otras muchas, se la identificó en una prospección del cielo realizada por el Ohio State University Radio Observatory, a una frecuencia de 1415 megahertzios. OJ287 se manifiesta absolutamente estelar; no ha podido detectarse ninguna nebulosa circundante ni siquiera en las fotografías de mayor exposición. Sólo una minoría de lacertae tienen halos discernibles, por lo que se hace difícil distinguir a los objetos sobre placas fotográficas. La identificación de OJ287, sin embargo, no admitió dudas; tenía todas las propiedades que hacen de BL Lacertae algo muy interesante y notable y que han llegado a ser las características definitorias de la clase: un espectro de radio invertido, variabilidad en todas las longitudes de onda, fuerte polarización óptica y espectro óptico sin rasgos específicos.

No se puede llegar al conocimiento de los objetos BL Lacertae sin poseer una idea aproximada de a qué distancia se hallan. Si se trata de objetos similares a las estrellas instalados en nuestra propia galaxia, entonces su luminosidad intrínseca debe ser alta sin salirse de lo normal; e incluso podrían radiar 10 veces la energía del Sol. Ahora bien, si están a distancias cosmológicas, de modo que su luz deba cruzar una fracción importante del universo visible hasta llegar a nosotros, la cantidad de energía que se disipe es de un orden de magnitud totalmente diferente. Los lacertae se cuentan entre las fuentes de radiación del universo más potentes, emitiendo quizá el céntuplo del total de nuestra galaxia entera. (La galaxia radia aproximadamente 10^{10} veces más energía que el Sol.)

El debate sobre las distancias de los lacertae, ahora decidido en favor de la escala cosmológica, ha seguido un camino similar al debate sobre las distancias de los quasars. Los resultados son los mismos. Toda estimación de distancias cosmológicas descansa sobre la hipótesis de que el universo está expansionándose uniformemente. Observamos esta



SE EMITE RADIACION cuando un electrón u otra carga eléctrica es acelerada; la naturaleza de la aceleración determina la longitud de onda y otras propiedades de la radiación. La radiación térmica *(izquierda)* resulta de colisiones y excitaciones de átomos y electrones en el interior de un medio caliente y denso, verbigracia, la atmósfera de una estrella. Los sucesos que dan lugar a la radiación son aleatorios, obedecen a leyes estadísticas que confieren al espectro una forma bien definida y, para cualquier temperatura dada, una banda de longitudes de onda relativamente estrecha. La polarización de las ondas, o, en otras palabras, su orientación si pudiera ser vista de frente, permanece aleatoria. Se considera que la radiación de los quasars y de los objetos BL Lacertae proviene de otro proceso, el mecanismo de sincrotrón *(derecha)*. Se emite radiación de sincrotrón cuando se acelera continuamente a los electrones, obligándoles a seguir una órbita circular; en los objetos astronómicos, las órbitas son realmente helicoidales, girando alrededor de las líneas de flujo de un campo magnético. Las ondas electromagnéticas se emiten tangentes a las órbitas de los electrones con una gama amplia de frecuencias. La radiación de sincrotón está fuertemente polarizada, y las ondas se alinean paralelas al plano del movimiento circular.

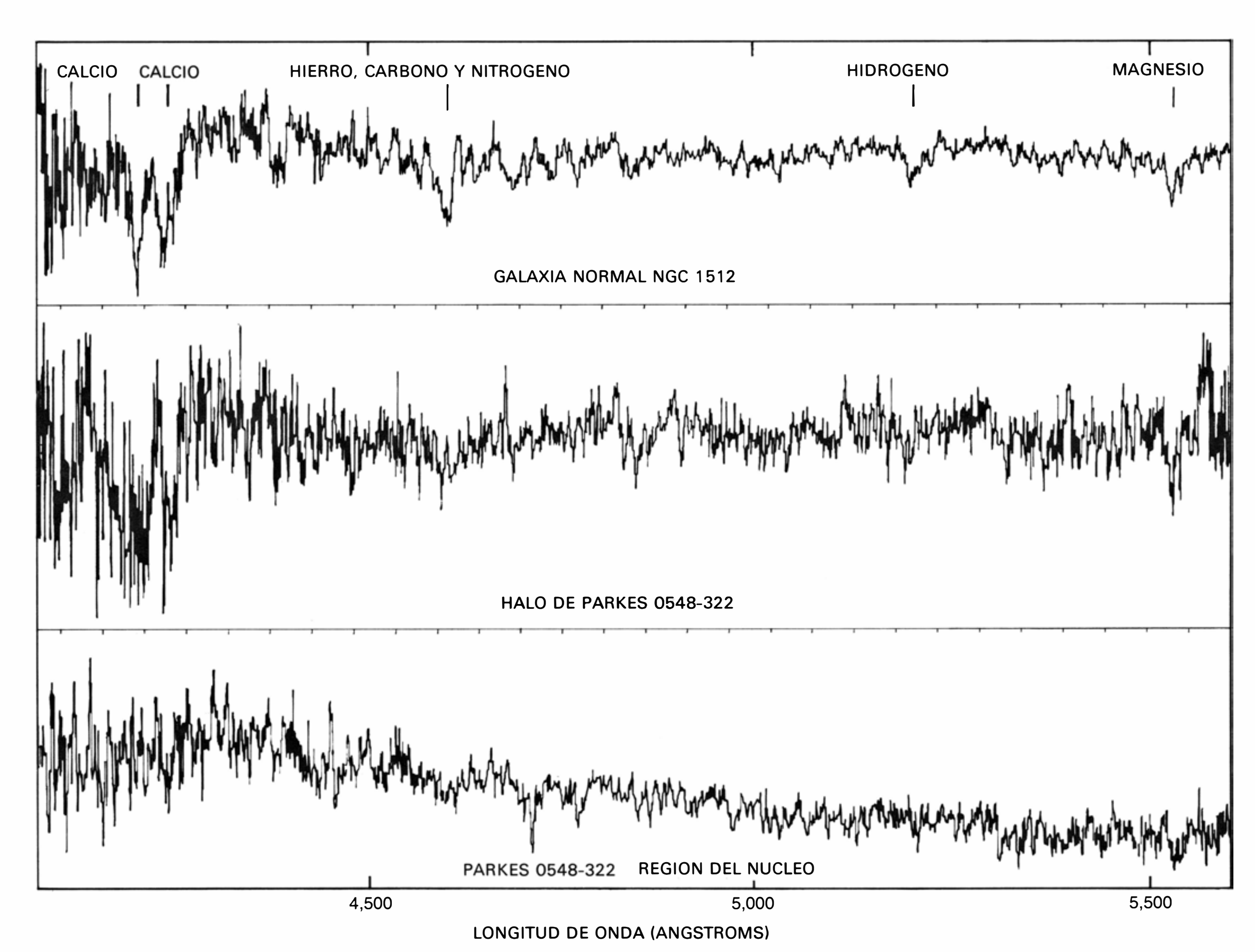
expansión en el movimiento de las galaxias lejanas, que parecen estar alejándose de nosotros y, a su vez, unas de otras. Como Edwin P. Hubble y Milton L. Humason demostraron hace casi 50 años, parece que la velocidad de recesión es proporcional a la distancia de las galaxias. A causa de ese movimiento, las características que una galaxia manifiesta en el espectro, como las líneas de absorción o de emisión, están desplazadas hacia longitudes de onda más largas; por tanto, las longitudes de onda en el rango visible están corridas hacia el extremo rojo del espectro. A partir de este corrimiento hacia el rojo puede determinarse la velocidad de recesión y, a partir de esta velocidad, deducirse la distancia.

Entre las propiedades específicas de los quasars se numera su acentuado corrimiento hacia el rojo. Si se atribuyen tales desplazamientos hacia el rojo a expansiones cosmológicas, entonces algunos quasars se están alejando de nosotros a una velocidad superior a la mitad de la velocidad de la luz y deben ubicarse a distancias inmensas. Nosotros podemos verlos porque son cientos de veces más luminosos que las mayores galaxias normales. Además, la luminosidad debe proceder de un volumen bastante pequeño. Se ha observado que algunos quasars varían su brillo a una escala temporal de pocos días, lo que pone de manifiesto que la región luminosa apenas alcanza escasos días luz de diámetro. De una región que es la millonésima parte de su diámetro, debe extraerse cien veces la potencia de una galaxia.

Como cabría esperar, algunos astrónomos han encontrado estas ideas difíciles de sostener y han propuesto otras alternativas. En las hipótesis de los quasars "locales", el problema de las enormes densidades de energía se orilla mediante la suposición de que los quasars están mucho más cerca y, por tanto, que son menos luminosos. Si bien entonces resulta inexplicable el fuerte corrimiento hacia el rojo que registran.

El debate ha seguido abierto a lo largo de doce años. La prueba crucial necesaria para corroborar la hipótesis cosmológica sería la demostración de que un quasar, al menos, está situado en una galaxia lejana y que las estrellas circundantes tienen el mismo corrimiento al rojo que el quasar. Sin embargo, no se ha observado ningún quasar con una nube definida de estrellas, ni hay por qué sorprenderse de no haberlo visto. La hipótesis cosmológica presupone una luminosidad para el núcleo tan alta que debe anular la contribución de las estrellas de fondo.

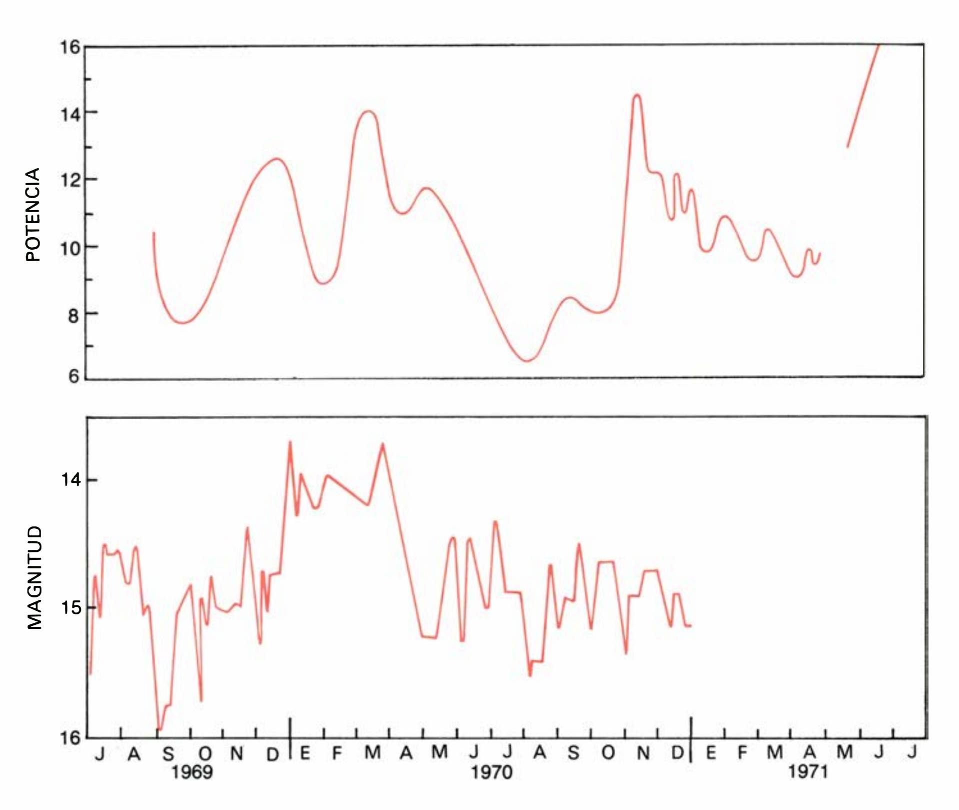
La necesidad de una medición que



LAS LINEAS DE ABSORCION en el espectro del halo borroso que rodea al objeto BL Lacertae apoyan la suposición de que se trata de una galaxia independiente situada a una gran distancia de nosotros. El objeto BL Lacertae está representado por Parkes 0548-322; se compara su espectro con el de una galaxia normal, la NGC 1512. Las líneas de absorción se impresionan sobre el espectro por los átomos de las capas más exteriores de las estrellas, proporcionalmente frías; su presencia en el espectro puede considerarse como prueba de que el halo está formado por estrellas. La ausencia de estas líneas en un espectro que registre sólo el núcleo del objeto es coherente con la hipótesis de que la luz se emite en esa región por una fuente no térmica. Por comparación, las líneas de absorción del objeto BL Lacertae se muestran en coincidencia con las de NGC 1512; naturalmente, se hallan desplazadas hacia las longitudes de onda más largas por la expansión general del universo, que origina que el Parkes 0548-322 se vaya alejando de nuestra galaxia a una velocidad de 21.000 kilómetros por segundo.

corrobore el corrimiento hacia el rojo destaca la importancia de los bordes borrosos vistos alrededor de algunos objetos BL Lacertae. Si estos bordes son estrellas, entonces el corrimiento hacia el rojo del objeto entero puede determinarse a partir de las líneas de su espectro. Por paradójico que parezca, este método ofrece la única esperanza de calibrar las distancias a que se encuentran los lacertae, ya que el espectro del núcleo luminoso no tiene ni líneas de emisión ni líneas de absorción para reconocerlo. El reto planteado a los observadores estriba en registrar el espectro de una nebulosa débil en la inmediata vecindad de su núcleo brillante y estrellado. El trabajo exige la inversión de mucho tiempo de telescopio, unas condiciones atmosféricas ideales y buena suerte en detectar los núcleos luminosos en intensidad mínima.

En 1974, dos astrónomos del observatorio de París mostraron que la región nebular que rodea a BL Lacertae tiene el tamaño aparente y la distribución de brillo propios de una galaxia elíptica gigante que se halle a una distancia correspondiente a una velocidad recesional de 21.000 kilómetros por segundo, el 7 por ciento de la velocidad de la luz, aproximadamente. Poco tiempo después, otros astrónomos de los Hale Observatories de California publicaban las mediciones obtenidas del desplazamiento al rojo de esta región nebular. Consiguieron sus resultados mediante ocultaciones mecánicas del núcleo, recibiendo sobre el espectrógrafo sólo la luz que provenía del halo. A partir de sus datos calcularon un corrimiento hacia el rojo que era justamente igual al propuesto por los astrónomos franceses. La prueba de este valor del corrimiento hacia el rojo no resultaba, sin embargo, convincente: consistía sólo en una pequeña depresión en el espectro registrado. Para mayor frustración, otro grupo de astrónomos del Lick Observatory fallaron en su intento de reproducir la observación. La validez de los resultados permanece dudosa, pero las investigaciones actuales siguen su curso.



EL REGISTRO DE LAS VARIACIONES de intensidad de radiación del BL Lacertae tiene una relación directa con el tamaño de la región radiante del núcleo. Las variaciones de radio *(arriba)* quedaron registradas durante dos años en una longitud de onda de 2,8 centímetros. Las fluctuaciones mayores tienen una escala de tiempo de pocas semanas. En las longitudes de onda ópticas *(abajo)* se observaron grandes cambios de magnitud en periodos correspondientes a varios días. Ningún objeto puede cambiar su luminosidad a una velocidad mayor que el periodo requerido por una señal para cruzar dicho objeto a la velocidad de la luz. En consecuencia, las emisiones ópticas de BL Lacertae deben provenir de una región de escasos días luz de diámetro; la región de radioemisión no puede ser mayor. En comparación, el sistema solar tiene un diámetro de medio día luz.

Mientras tanto, otros astrónomos del Mount Stromlo Observatory, en Australia, trabajaban en el registro de una línea espectral de AP Librae, un objeto localizado en el hemisferio Sur que guarda un notable parecido con BL Lacertae. No observaron la luz del núcleo galáctico, sino que, durante tres años, obtuvieron cierto número de espectros, a la espera de detectar el núcleo en el mínimo de luminosidad. Ningún espectro, por sí solo, proporcionó pruebas evidentes, pero tomados juntos los espectros resultaron convincentes. Aparecían varias líneas débiles de absorción, que suelen darse en el espectro de las galaxias. El corrimiento hacia el rojo de estas líneas era equivalente a una velocidad de recesión de 15.000 kilómetros por segundo y a una distancia del orden de mil millones de años luz. Por tanto, AP Librae es definitivamente un objeto extragaláctico. Al parecer, consta de una pequeña región de excepcional luminosidad, situada en el núcleo de una extensa galaxia elíptica.

Los corrimientos hacia el rojo quedaron determinados también con razonable confianza para cuatro lacertae más. Todos parecen hallarse en galaxias a gran distancia. Corrimientos hacia el rojo de 15.000 a 20.000 kilómetros por segundo no son parangonables naturalmente con los desplazamientos extremos al rojo de los quasars más distantes. No debe concluirse, sin embargo, que los quasars sean los objetos más distantes. Las diferencias en las mediciones respecto al corrimiento hacia el rojo pueden resultar de observaciones sesgadas: puesto que los espectros de lacertae son tan débiles, es probable que sólo se hayan registrado los más próximos. Aun así, las líneas de absorción provisionalmente identificadas en el espectro de dos lacertae sugieren que se mueven a una velocidad que es aproximadamente la mitad de la velocidad de la luz. Uno de los objetos, designado por AO 0235+164, tiene dos conjuntos de líneas con diferentes corrimientos hacia el rojo, equivalentes al 52 por ciento y al 85 por ciento de la velocidad de la luz. Un espectro dual de este tipo puede producirse por absorción de dos galaxias que se interpongan en la línea de luz, lo cual requeriría que el objeto se hallara más alejado que cualquiera de las dos galaxias. (Otra interpretación menos probable es que las líneas sean producidas por envolturas de gas escapadas de una galaxia a muy alta velocidad.) A la distancia indicada por el corrimiento hacia el rojo, la galaxia no podría registrarse y sólo el núcleo, parecido a una estrella, sería detectado.

Si los objetos BL Lacertae están a distancias cosmológicas ¿qué es lo que origina su luminosidad deslumbrante? La fuente fundamental de energía se desconoce en absoluto, aunque en los últimos años ha vuelto a prosperar la idea que la hace gravitacional. Existe, además,

un mecanismo conocido que puede aumentar en su etapa final y producir gran cantidad de radiación. Este mecanismo es el proceso de sincrotrón, interacción de electrones de alta energía con un campo magnético.

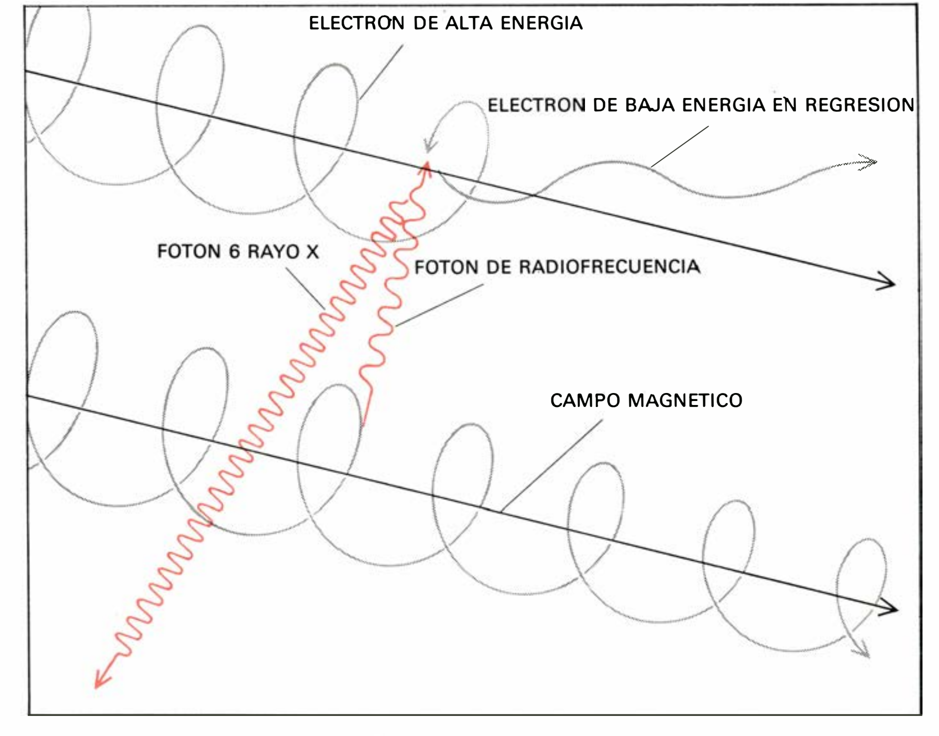
Se emite radiación electromagnética cuando un electrón es acelerado, esto es, forzado a cambiar su velocidad o su dirección. Se emite radiación de sincrotrón cuando un electrón, que se mueve a una velocidad próxima a la de la luz, sufre una aceleración continua obligándole a seguir una trayectoria circular. En condiciones terrestres, ello puede lograrse cuando los electrones son acelerados a velocidades relativistas (próximas a la velocidad de la luz) en un acelerador circular de partículas, como puede serlo el sincrotrón, y de ahí el nombre de la radiación. En un contexto astronómico, la radiación de sincrotrón se produce cuando una población de electrones relativistas encuentra un campo magnético. Los electrones son desviados por el campo según trayectorias helicoidales orientadas de acuerdo con las líneas del flujo magnético. La radiación es emitida tangente a estas trayectorias, con una distribución de frecuencia determinada por la intensidad del campo y la energía de los electrones.

La radiación de sincrotrón se distingue claramente de la radiación térmica de las estrellas ordinarias. La distribución espectral de la radiación de sincrotrón es mucho más ancha; en virtud de lo cual, los objetos astronómicos que emiten por el mecanismo de sincrotrón pueden ser captados en ambas longitudes de onda de radio y ópticas, mientras que la radiación térmica de las estrellas está fuertemente confinada a una banda única de longitudes de onda. Otro hecho que define a la radiación de sincrotrón es su polarización: las ondas electromagnéticas tienden a alinearse paralelas al plano del movimiento circular del electrón emisor. La radiación térmica de las estrellas no es polarizada, es decir, las ondas están orientadas aleatoriamente. A su polarización se debe la prueba más convincente de la emisión de sincrotrón en los quasars y en los objetos BL Lacertae. En estos últimos, la polarización para longitudes de onda visibles es típicamente del 10 por ciento, y puede llegar al 35 por ciento. De ello se sigue que los campos magnéticos responsables de las emisiones de sincrotón tendrán una estructura ordenada.

La propiedad de los objetos BL Lacertae que cuesta entender más es el tamaño reducido de la región emisora de la radiación de sincrotrón. El tamaño se infiere de la razón con que se ve variar la luminosidad de los objetos. Los tamaños máximos son ya peligrosamente pequeños para los quasars, y el problema se agrava en el caso de los lacertae.

Habrá que disponer de otras muchas observaciones para definir el rango de variabilidad que poseen los lacertae, pero se pueden avanzar ya algunas estimaciones. El propio BL Lacertae ha sido el más cuidadosamente registrado para cambios rápidos. Para longitudes de onda ópticas, el rango de intensidades registrado cubre alrededor de tres magnitudes, o un factor de 15. La luminosidad ha cambiado como mucho un 400 por ciento en dos días; en un intervalo de minutos se ha registrado una oscilación de unos pocos por ciento. La polarización también ha variado en grados y en orientación, a lo largo de escasas noches. Por lo que se refiere a ondas de radio, la variación es menos extrema: aunque el rango de luminosidades puede alcanzar el 300 por ciento, se toma la semana como escala de tiempo mejor que el día, lo que sugiere que las emisiones de radio provienen de un volumen mayor que las emisiones ópticas. En BL Lacertae, las variaciones ópticas y de radio parecen totalmente independientes; solamente en uno de los lacertae, OJ287, se ha descubierto una correlación entre las dos partes del espectro. En los demás lacertae, y a todas las longitudes de onda, las variaciones parecen ser erráticas; no se han percibido ni trayectorias ni periodos. Se han registrado algunas erupciones espectaculares. En 1975, el objeto AO 0235 + 164 aumentó su brillo en un factor de 100; otro lacertae, PKS 2155-152, incrementó su brillo óptico seiscientas veces.

La rápida variación óptica en los objetos BL Lacertae sugiere que la región del núcleo tiene un diámetro de sólo unos pocos días-luz, y podría ser sustancialmente menor. (El sistema solar, si se define por la órbita de Plutón, tiene un diámetro de aproximadamente medio día-luz). Las variaciones en los lacertae pueden ser diez veces más rápidas que las de los quasars, por tanto cabe que el núcleo luminoso de un objeto BL Lacertae sea la décima parte del de un quasar. En medidas de volumen, esta razón debe ser cubicada y, así, los lacertae



LA DISPERSION COMPTON INVERSA es un proceso que, se creería, prohíbe regiones de radiación de radiofrecuencias extraordinariamente intensas, tales como las observadas en los objetos BL Lacertae. La dispersión tiene lugar cuando un fotón, o cuanto de energía electromagnética, colisiona con un electrón de alta energía. Parte de la energía del electrón se transfiere al fotón, el cual es convertido de este modo a una longitud de onda mucho más corta, de ordinario en la gama de los rayos X. La dispersión Compton es un efecto poco común, ya que las colisiones entre fotones y electrones de alta energía son improbables. En el núcleo de los objetos BL Lacertae, sin embargo, la densidad tanto de fotones como de electrones de alta energía debe ser muy alta. Cálculos basados sobre el tamaño aparente y la luminosidad del núcleo predicen una "catástrofe Compton inversa", en la que toda radiación se manifiesta como una descarga de rayos X, la cual rápidamente desposee a los electrones de su energía, según se explica en el texto.

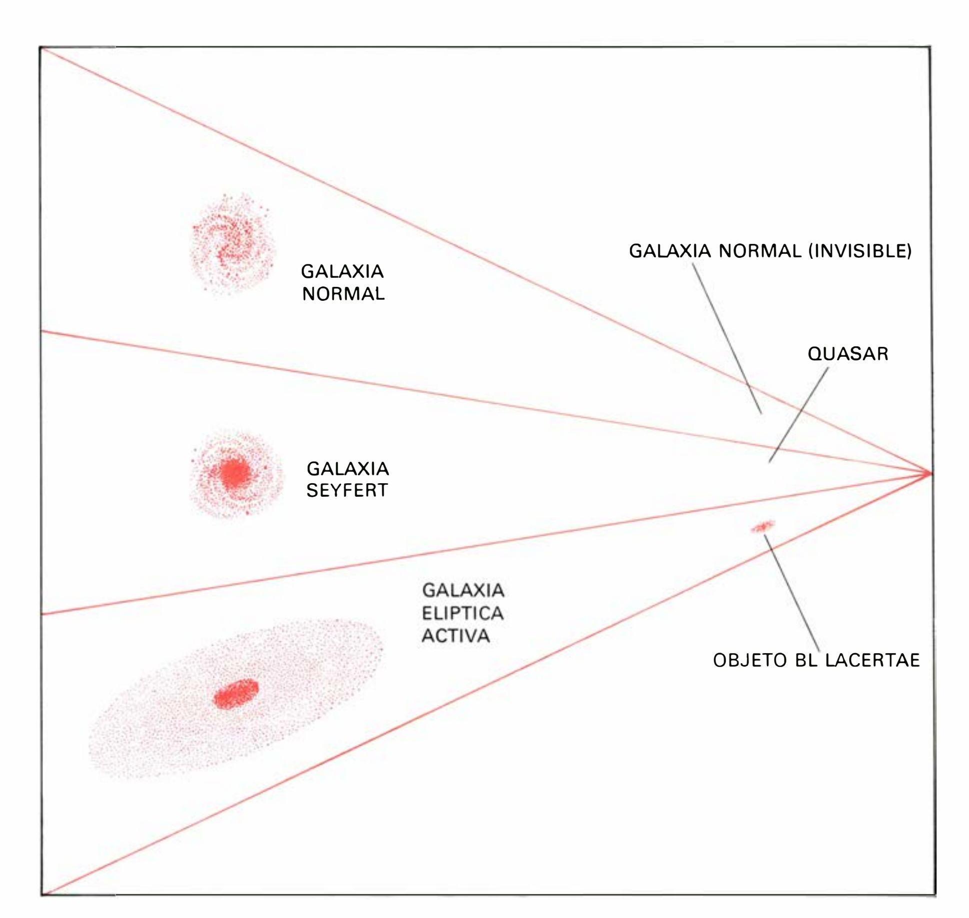
pueden radiar la misma energía que un quasar con una milésima parte del volumen.

Por física teórica se sabe que la radiación de tanta energía, a partir de volumen tan pequeño, por emisión de sincrotrón, no puede mantenerse mucho tiempo. La radiación se extinguiría, al parecer, por el llamado efecto de catástrofe inversa de Compton.

El efecto Compton recibe tal nombre en honor del físico americano Arthur Holly Compton, quien fue el primero en explicarlo en 1922. Describe la dispersión de un fotón, o cuanto de energía electromagnética, a partir de un electrón dotado de una pequeña energía cinética. El fotón cede parte de su energía al electrón, transición que se contempla como una ligera disminución en la frecuencia (o aumento en la longitud de onda) de la radiación. Se observa la dispersión Compton inversa cuando un fotón encuentra un electrón con alta energía. Entonces es el electrón quien cede energía al fotón. Más aún, la variación de la frecuencia no es mayor que la débil variación anterior: un fotón de radiofrecuencia puede ser convertido por la dispersión Compton inversa en un rayo X de alta energía.

La probabilidad de que un fotón colisione con un electrón relativista suele ser muy pequeña y, un suceso raro, la dispersión Compton inversa. No obstante, la probabilidad de colisión aumenta cuando electrones y fotones están confinados a alta densidad. A la densidad requerida para explicar el tamaño y luminosidad de objetos BL Lacertae, todos los fotones emitidos por el proceso de sincrotrón serían dispersados por electrones y convertidos en rayos X.

El efecto Compton inverso puede describirse de una forma equivalente imaginando la interacción desde el punto de vista de un electrón. En el núcleo de un objeto BL Lacertae, un electrón está sujeto a dos campos magnéticos. Uno es el campo externo, original, que confina los electrones en trayectorias helicoidales y da lugar a la radiación de sincrotrón. El otro es el campo de la radiación misma, el campo magnético, que es una componente de cada onda electromagnética. A baja densidad, el campo de radiación es despreciable, pero cuando la densidad de fotones aumenta resulta ser el más fuerte de los dos campos. Los electrones responden luego principalmente al campo de radiación; a consecuencia de lo cual las radioondas se suprimen y en su lugar se emiten rayos X. Los electrones son rápidamente frenados y cesan de radiar de inmediato.



LA HIPOTESIS DEL NUCLEO ACTIVO nos da una única explicación posible de ciertas galaxias poco comunes que están relativamente próximas, de los quasars y de los objetos BL Lacertae. Todos los objetos tienen en común la emisión de intensa radiación con un espectro no térmico; por otra parte, la variabilidad de la radiación sugiere que, en muchos casos, debe venir de una región del núcleo bastante pequeña. Las galaxias espirales que tienen tales núcleos son llamadas galaxias Seyfert; si se las confinara sumamente distantes se parecerían mucho a los quasars, ya que sólo podría verse el núcleo brillante y no las estrellas que lo rodean. A la misma distancia, una galaxia normal resultaría demasiado débil como para ser detectada. Los núcleos activos se encuentran también en algunas galaxias elípticas y, si las viéramos a una gran distancia, tendrían la misma apariencia que los objetos BL Lacertae. Débiles halos de estrellas en torno al núcleo sólo serían visibles en los lacertae y en los quasars que se encuentran más próximos a nuestra galaxia.

Estos razonamientos apoyan la idea de que los objetos BL Lacertae no pueden estar a muy grandes distancias; si así fuera emitirían rayos X, los cuales no son observados, en vez de radiondas, que sí lo son. Obviamente, este razonamiento contradice algunas de las pruebas observacionales. Se ha propuesto un número de medidas *ad hoc* para resolver la contradición, pero ninguna de ellas resulta convincente. No se puede esperar el éxito de los razonamientos teóricos cuando no se comprende bien la física.

Con el descubrimiento de numerosas galaxias activas, quasars y objetos BL Lacertae, se ha reconocido que estas categorías se mezclan suavemente unas con otras. Los quasars más variables, por ejemplo, pueden alterar su luminosidad en 100 por ciento en pocos días, aproximándose a la variación que sufren los lacertae. Los mismos quasars tienden a tener proporcionalmente espectro rojo y alta polarización óptica, que son característicos de los lacertae. Análogamente, algunas radiogalaxias, dotadas de regiones brillantes nucleares manifiestan un estrecho parecido con los objetos BL Lacertae, aunque tienen líneas de emisión en su espectro. Las líneas son del mismo tipo, pero no de la misma intensidad, que las líneas encontradas en un espectro de quasar. Algunas galaxias elípticas cercanas carecen de núcleo brillante en longitudes de onda visibles, pero lo tienen en longitudes de onda de radio; más aún, la radiación visible de estas galaxias está a veces ligeramente polarizada, lo cual sugiere que una débil región de radiación óptica no térmica puede estar por debajo de la radiación más potente de estrellas normales. En un caso dramático, una radiogalaxia, 0521-36, con líneas de emisión en longitudes de onda ópticas, brillaba hasta resultar un objeto BL Lacertae con un espectro sin rasgos distintivos.

La existencia de formas intermedias refuerza la conjetura de que las galaxias activas, los lacertae y los quasars son manifestaciones de un mismo y único proceso. Cabe la posibilidad de que todos estos objetos sean galaxias con regiones compactas en sus núcleos, que emiten radiación intensa y polarizada a lo largo de un ancho espectro. La energía salida de estas pequeñas regiones luminosas varía en una amplia gama, de galaxia a galaxia y de tiempo en tiempo en una misma galaxia. El núcleo más potente puede ser cientos y miles de veces más brillante que el resto de la galaxia; por cuya razón, el núcleo estrellado solamente puede detectarse en los quasars y en algunos lacertae, como el OJ287. Cuando el núcleo y la galaxia madre tienen aproximadamente el mismo brillo, los vemos ambos, como en AP Librae y quizás en el mismo BL Lacertae. Cuando el núcleo es débil, su radiación queda eclipsada por la radiación de las estrellas ordinarias y la galaxia aparece como normal, aunque las emisiones de radio puedan detectar incluso la presencia de un núcleo activo.

Una posible objeción a este esquema unificado es que todos los lacertae conocidos son radiofuentes, mientras que muchos quasars no tienen emisiones de radio detectables. En efecto, la mayoría de los quasars pueden ser radioemisores en reposo, tal vez incluso en la proporción del 95 por ciento. Esta distinción, no obstante, puede reflejar simplemente un sesgo instrumental. Los primeros quasars fueron detectados por su brillo poco común a longitudes de onda de radio; naturalmente, todos ellos eran radiofuentes. Sólo después, cuando se acometió la investigación del cielo en longitudes de onda ópticas para objetos brillantes en el azul y el ultravioleta, se descubrieron los quasars radiosilenciosos. Puede concluirse que los lacertae radiosilenciosos abundan también, pero debido a que sus aspectos ópticos son menos diferenciados que los de los quasars, no se ha encontrado todavía ninguno de ellos.

La diferencia más intrigante entre quasars y lacertae estriba en que aquéllos poseen intensas líneas de emisión en su espectro, de las que carecen los lacertae. Se desconoce por qué ocurre así.

Las líneas de emisión se producen cuando se calienta un gas a unos 10.000 grados Kelvin; a esta temperatura, muchos átomos pierden electrones, y se dice que el gas está ionizado. Cuando un electrón y un ion se recombinan, el electrón desciende una serie de niveles cuánticos, perdiendo energía en cada

transición; la energía aparece en forma de radiación electromagnética con longitudes de onda características de cada especie atómica. Al ser el hidrógeno el elemento más común, las líneas de recombinación del hidrógeno resultan generalmente las más intensas.

Se supone que, en los quasars, las líneas de emisión se producen cuando los gases son calentados por radiación ultravioleta procedente del núcleo luminoso. Los fotones ultravioleta tienen suficiente energía para ionizar los átomos de hidrógeno (y otros átomos también), y son fuertemente absorbidos. La energía absorbida por el gas de esta forma podría calentarlo y permitir una rápida expansión de la nube. A su vez, la expansión puede explicar por qué las líneas de emisión son bastante anchas: fracciones de una nube en expansión son empujadas hacia el observador y otras a otro sitio, dando lugar a corrimientos opuestos en la longitud de onda.

La ausencia de líneas de emisión intensas en objetos BL Lacertae permanece confusa, si bien se han avanzado, por lo menos, dos explicaciones prometedoras. Una posibilidad es que los lacertae emitan muy pocas radiaciones ultravioleta como para que ionicen completamente una nube de gas. El espectro de la mayoría de los lacertae disminuye rápidamente en la banda visible, hasta el punto de que parece que no tiene tanto brillo en la zona ultravioleta como los quasars. Por otra parte, algunos lacertae, verbigracia el OJ287, tienen un espectro óptico configurado como el de los quasars, aunque faltan todavía las líneas de emisión. La distribución puede ser establecida registrando el espectro ultravioleta de un objeto BL Lacertae, aunque las longitudes de onda en cuestión son fuertemente absorbidas por la atmósfera terrestre.

Una explicación más simple de la ausencia de líneas de emisión es que las galaxias que envuelven los lacertae no contienen gas que se ionice por las emisiones ultravioleta del núcleo. La plausibilidad de esta hipótesis queda corroborada por la observación ya que, cuando una galaxia puede resolverse rodeando un objeto BL Lacertae, se trata siempre de una galaxia elíptica. Estudios de galaxias elípticas cercanas han demostrado que, consideradas en grupo, son deficientes en gas. Por otra parte, núcleos activos enclavados en galaxias espirales poseen líneas de emisión intensas, como ocurre con las que definen a los quasars. En galaxias espirales, la nuestra por ejemplo, el gas abunda.

A partir de esta observación podemos esbozar un esquema que nos dé una síntesis en que se relacionen estos objetos raros. Los quasars resultan núcleos activos en galaxias espirales que han sido desplazadas a una gran distancia, en donde su nebulosidad galáctica está completamente oscurecida. A distancias menores, el mismo objeto se llamará galaxia Seyfert. Cuando el mismo tipo de núcleo activo se observa en una galaxia elíptica a gran distancia podemos identificarlo como objeto BL Lacertae.

Es una síntesis sugestiva que no está exenta de dificultades. Por ejemplo, algunas galaxias elípticas tienen líneas de emisión intensas, y ciertos quasars poseen una estructura de radio de dos lóbulos que no se ha encontrado en galaxias espirales. Ambas observaciones tienden a asociar los quasars con las galaxias elípticas, mejor que con las espirales.

En el fondo, los grandes interrogantes siguen sin respuesta. ¿Cuál es la fuente de energía de las galaxias activas y de la mayoría de los objetos distantes? Ya que ningún mecanismo conocido por los físicos terrestres lo explican, la atención se ha dirigido hacia propuestas exóticas y especulativas. Durante cierto tiempo, tuvo sus adictos la hipótesis de la superestrella, una enorme bola de gas caliente con una masa millones de veces mayor que la del Sol. Más recientemente, la devoción popular se ha desplazado hacia los agujeros negros, el objeto compacto que es el último residuo de materia después del colapso gravitacional. Así, hasta un rosario de posibles fuentes: "agujeros blancos", núcleos retardados, púlsares gigantes, colisión de estrellas, galaxias colapsadas y supernovas múltiples.

Un modelo convincente deberá explicar el hecho de que quizá se trate del mecanismo más potente del universo y deberá situarlo en un compartimiento de tamaño reducido. Todos los modelos se acrisolarán en la prueba crucial: la variación de luminosidad más rápida que pueda ajustarse, pues ese período corresponderá al tiempo requerido por una señal para atravesar la región emisora. Han tenido que descartarse algunos modelos porque no permitían variaciones suficientemente rápidas. El modelo de agujeros negros puede servir para alta luminosidad y para las variaciones más rápidas observadas hasta el momento en los lacertae, exigiéndoles que posean un diámetro de 3 millones de kilómetros (aproximadamente la mitad del tamaño del Sol). Un objeto de ese tamaño no puede fluctuar significativamente en un tiempo inferior a los diez segundos. Si observamos en un objeto BL Lacertae variaciones más rápidas, el modelo de agujeros negros deberá descartarse.

Comunicaciones ópticas

La primera prueba comercial de servicio telefónico óptico está realizándose en Chicago. Se envían las señales a través de fibras de vidrio en forma de impulsos generados por pequeñas fuentes de luz de estado sólido

W. S. Boyle

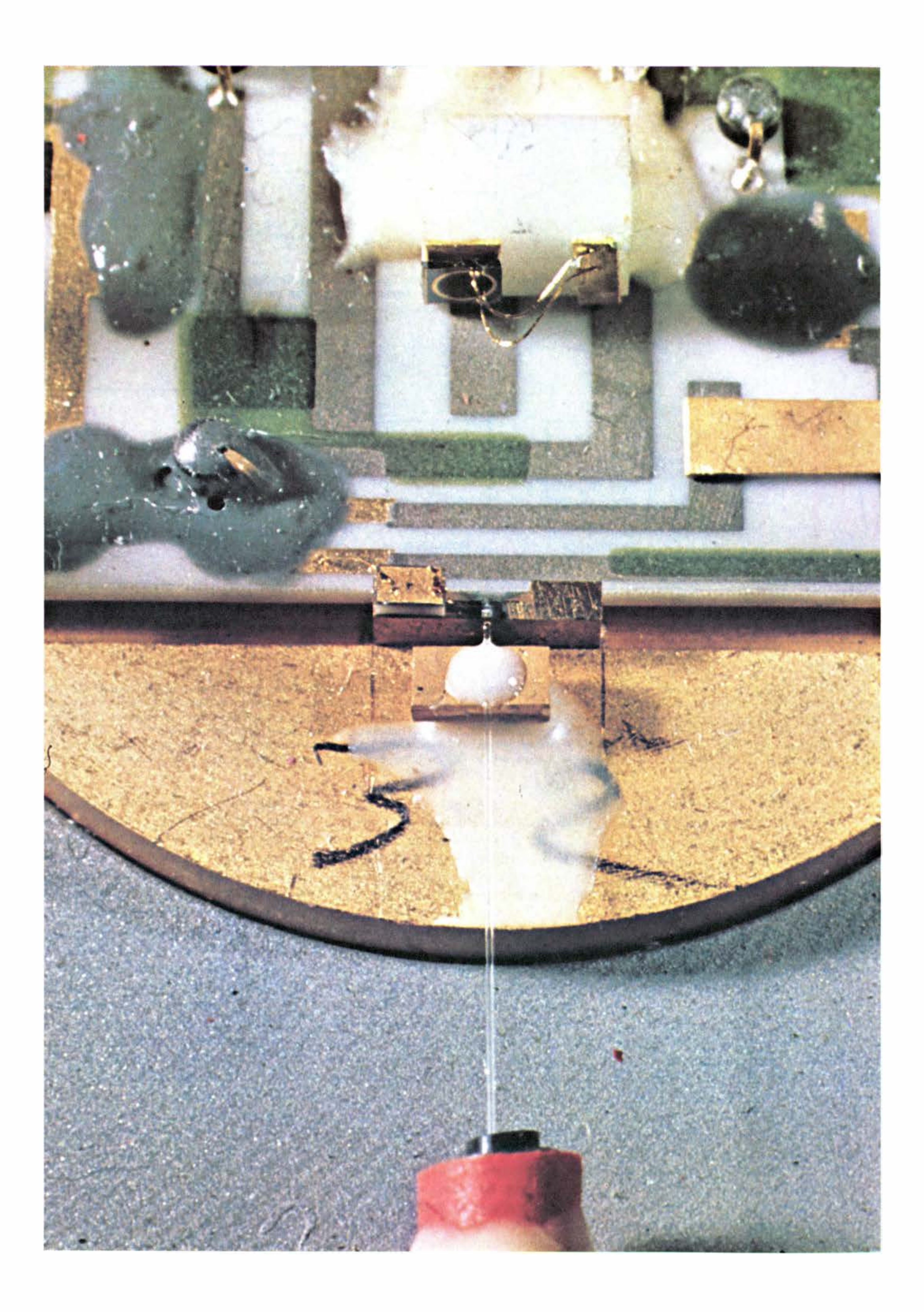
Hace tres meses, la compañía Bell System comenzó a ensayar la comercialización de un sistema óptico en el que los mensajes se codifican en impulsos de luz que se transmiten a través de fibras de vidrio del grosor de un cabello. El nuevo sistema transporta señales de voz, datos y vídeo a lo largo de unos dos kilómetros y medio de cable subterráneo, mediante la conexión de dos centrales telefónicas de la Illinois Bell Telephone Company y un gran edificio mercantil en el centro comercial de Chicago. El cable de fibra óptica, que posee un diámetro de 12,5 milímetros, contiene 24 fibras en dos cintas de doce fibras cada una. La capacidad de información de cada fibra es de 44,7 megabits por segundo, lo cual significa que la fuente de luz que alimenta a la fibra se enciende y apaga 44,7 millones de veces por segundo. A esta velocidad, cada fibra puede transportar 672 señales de audio unidireccionales; por consiguiente, las 24 fibras tienen una capacidad de 12×672, es decir, 8064 conversaciones bidireccionales. Para igualar esta capacidad con pares convencionales de hilos de cobre se requeriría un cable muchísimo mayor. Aparte de las ventajas de esta tecnología, el sistema ahorrará cobre e incrementará grandemente la capacidad de las canalizaciones existentes.

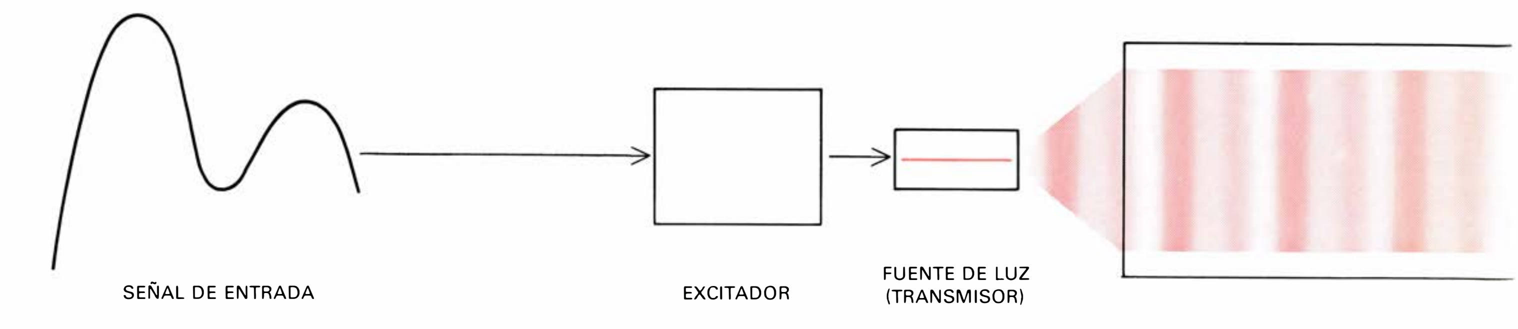
No hay nada particularmente nuevo en utilizar la luz para las comunicaciones. Después de todo, los indios americanos enviaban señales de humo y los ingleses encendían hogueras para avisar de la aproximación de la armada española. En 1790, Claude Chappe construyó un sistema telegráfico óptico formado por estaciones de semáforo sobre las cumbres de Francia. El sistema, que según la opinión general podía transmitir mensajes a distancias de 200 kilómetros en 15 minutos, permaneció en servicio hasta que fue sustituido por el telégrafo eléctrico. En 1880 Alexander Graham Bell inventó el "fotófono" con el que demostraba que la voz podía transmitirse a través de un rayo de luz. En el sistema Bell, se enfocaba un rayo estrecho de luz solar sobre un espejo delgado. Al tiempo que las ondas de sonido de la voz humana hacían que el espejo vibrase, variaba proporcionalmente la cantidad de energía luminosa transmitida a un detector de selenio. La luz recibida en el detector hacía variar a la resistencia del selenio, y, por tanto, la intensidad de la corriente en el receptor óptico. Y por lo menos hasta la Segunda Guerra Mundial, era común, entre los buques, intercambiar mensajes con señales de luz en código Morse.

La novedad reside hoy en las técnicas utilizables para generar un rayo de luz que pueda ser modulado a velocidades extremadamente altas, y para transmitir a varios kilómetros las señales resultantes a través de una fibra de vidrio, con una pérdida de energía aceptablemente baja. El interés por las comunicaciones ópticas data de la primera demostración del láser en 1960. Este dispositivo, capaz de emitir un rayo aproximadamente monocromático de intensa radiación visible o infrarroja, pertenece a una región del espectro electromagnético cuyas frecuencias eran 10.000 veces más altas que las frecuencias más elevadas, entonces en servicio, para sistemas de comunicación por radio. Puesto que la capacidad potencial de transporte de información aumenta directamente con la frecuencia, los ingenieros de comunicaciones emplearon durante décadas buena parte de su agudeza en el desarrollo de sistemas de frecuencias más altas. Desde los lejanos días del albor de la radio, se había ido avanzando hacia frecuencias gradualmente superiores hasta llegar a rozar los cinco órdenes de magnitud, de 100 kilohertzios (100.000 ciclos por segundo) a 10 gigahertzios (10^{10} ciclos por segundo). Ahora, el láser proporcionaba un incremento de cuatro órdenes de magnitud más, llegando a 100 terahertzios (100 billones de ciclos por segundo). Utilizando solamente una pequeña fracción de la gama de la frecuencia óptica generada por el láser, un simple sistema óptico podía, en principio, transportar simultáneamente las conversaciones telefónicas de todos los habitantes de América del Norte.

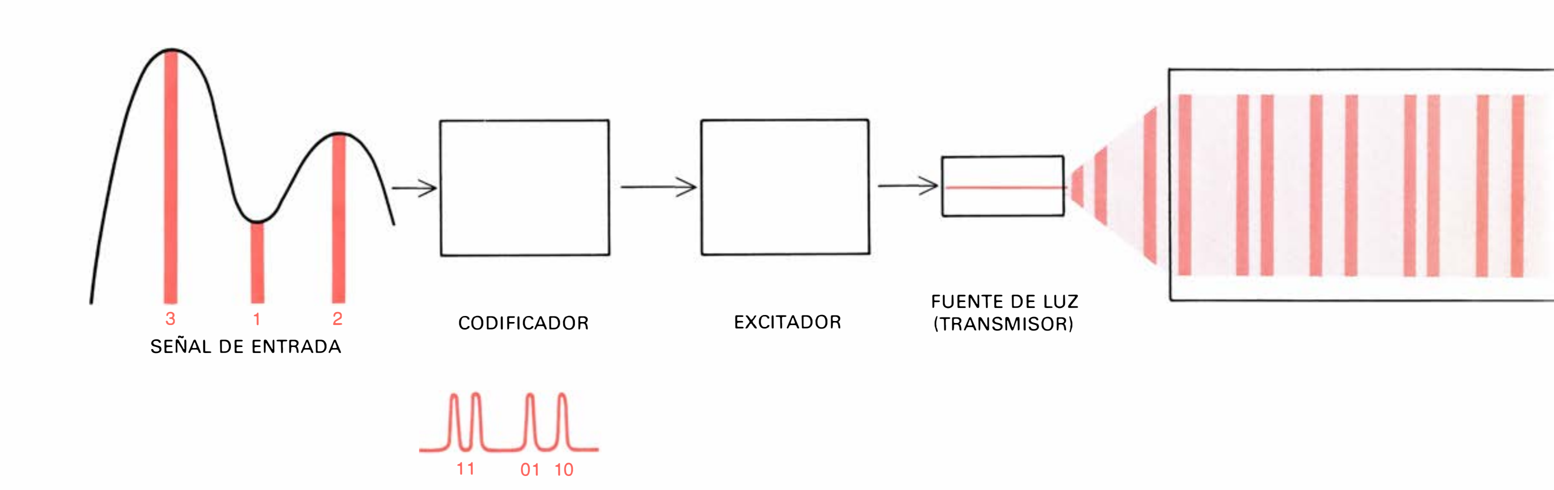
Los primitivos láseres, sin embargo, fueron voluminosos y poco fiables; el mejor de ellos fallaba a los pocos meses de funcionamiento. Posteriormente, se comprobó que la idea de enviar rayos láser de un sitio a otro, a través de la atmósfera, en analogía con el sistema de radioenlace por microondas, no resultaba satisfactorio porque la señal quedaría atenuada por la niebla, la polución, la lluvia y la nieve. En realidad, es más fácil transmitir pulsos de luz de Arizona a la Luna que transmitirlos de un extremo al otro de Manhattan.

EL CIRCUITO REGENERADOR OPTICO recibe los impulsos de luz debilitados, entregados por una fibra óptica de un sistema de comunicación óptica, y los amplifica en nuevos impulsos para continuar su propagación. Esta fotografía, realizada en los Bell Laboratories, reproduce con una ampliación de 10 veces una parte del circuito amplificador y de la fibra óptica que transportará los impulsos regenerados. La fibra *(abajo)* emerge del cable marcado en rojo. El extremo de la fibra está sujeto firmemente a un pequeño láser de arseniuro de galio mediante una gota de resina epoxy. Los impulsos de luz regenerados, producidos por el láser, pueden propagarse hasta 14 kilómetros a través de la fibra, antes de regenerarlos de nuevo. Un fotodiodo, encapsulado en el bloque blanco adyacente, recoge la luz emitida por la parte posterior del láser y la utiliza para compensar el efecto de las variaciones de temperatura sobre el circuito de conducción del láser.





SISTEMAS DE COMUNICACION OPTICA, que pueden diseñarse para transmitir señales codificadas de distintos modos. En el sistema "analógico" más simple, representado aquí, la amplitud de la señal de llegada queda convertida directamente en variaciones de amplitud del rayo luminoso que penetra en la fibra óptica. Un fotodetector, que se halla en el receptor, convierte la variación de intensidad de luz en la



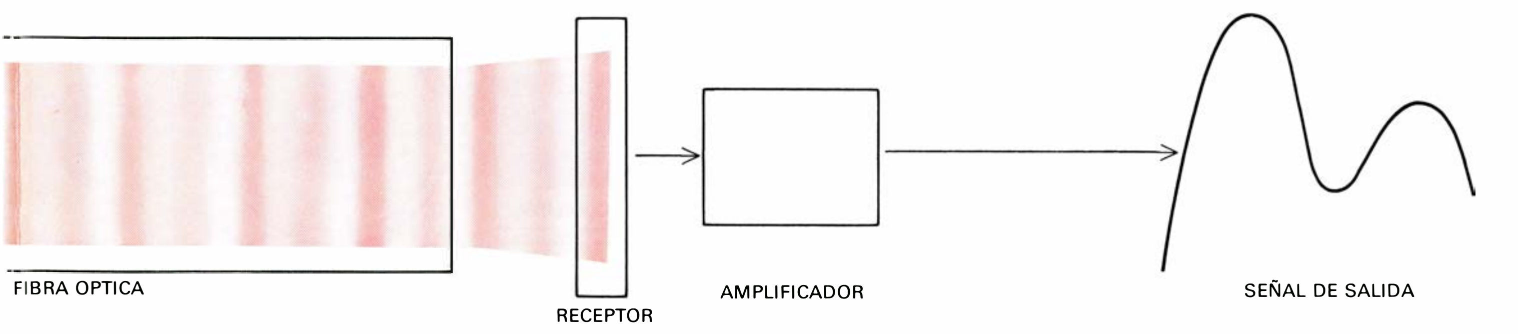
LA CODIFICACION DIGITAL está siendo adoptada para los sistemas de comunicación óptica comerciales, debido a sus muchas ventajas sobre la codificación analógica. En la codificación digital, la amplitud o altura de la onda de entrada ***(izquierda)*** **es muestreada electrónicamente a intervalos regulares** ***(barras verticales debajo de la onda de entrada).*** **Para una representación exacta, la onda debe ser muestreada a una velocidad doble de la frecuencia de su componente más alta. Así, una señal de audio, cuya frecuencia máxima sea de 4000 ciclos por segundo, deberá ser muestreada a 8000 veces por segundo. Las alturas de las muestras individuales se codifican en secuencias de dígitos binarios: ceros y unos. Para la transmisión, un uno puede representarse por un impulso y un cero por la ausencia del mismo. En un sistema vocal**

Se ha registrado un firme progreso en la realización de láseres compactos, fiables y de larga duración, así como en buscar alternativas para la transmisión de los rayos de luz a través de la atmósfera. Aunque para algunas aplicaciones se sigue prefiriendo el uso del láser, para otras resulta adecuado un dispositivo más simple y barato, el diodo emisor de luz de alta intensidad (LED). La primera alternativa prometedora para transmitir señales de luz a través de la atmósfera consistía en enviar señales ópticas por un conducto luminoso, que constaba de un tubo cuidadosamente fabricado, de un centímetro más o menos de diámetro, el cual proporcionaba con la ayuda de medios ópticos (posiblemente variaciones locales en la densidad del gas en el tubo) la manera de desviar los rayos cuando dicho tubo se separase de una línea recta.

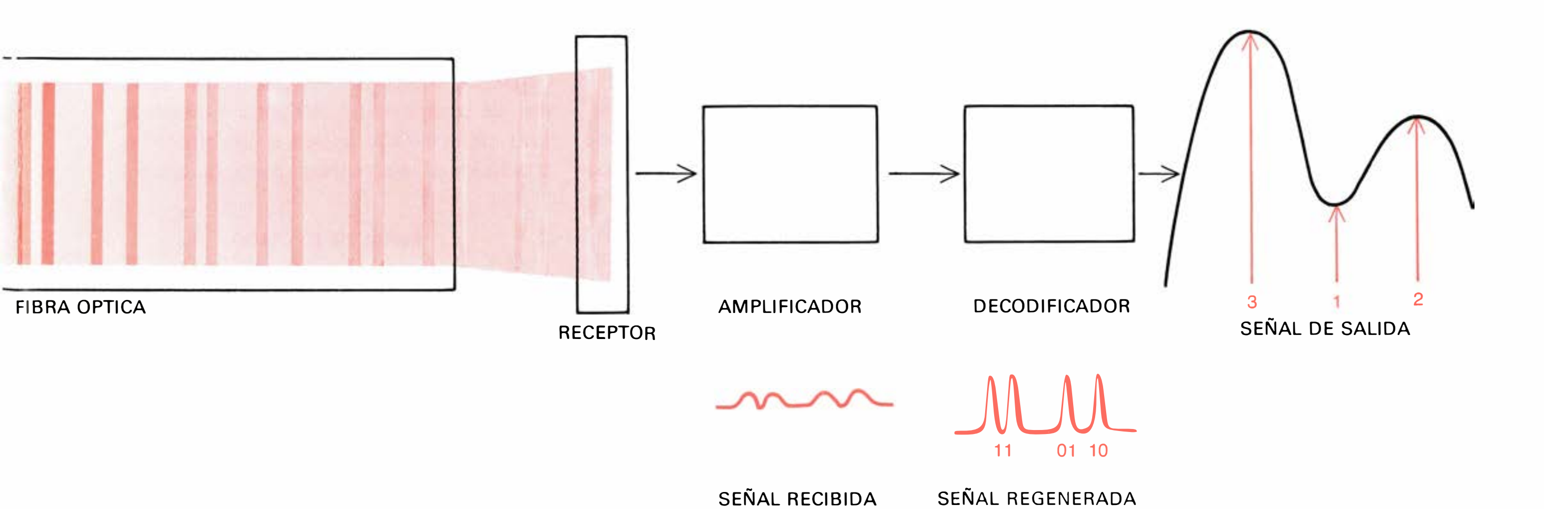
Como una alternativa al tubo de luz, que presentaba muchos problemas de orden práctico, los ingenieros de comunicaciones comenzaron a estudiar la posibilidad de transmitir luz a través de fibras de vidrio. Durante algún tiempo, se habían venido utilizando paquetes de fibras de vidrio o plástico para transportar luz a distancias cortas, por ejemplo, para iluminar un panel de instrumentos o para examinar el interior del estómago, pero no eran lo bastante transparentes como para utilizarlas en comunicaciones ópticas. Los materiales comúnmente empleados eran menos transparentes que el agua. Las fibras de vidrio últimamente desarrolladas para comunicaciones poseen una transparencia tal que si el agua del mar fuese tan clara como ellas, se podría ver el fondo del océano más profundo.

Antes de considerar qué tipos de sistemas de comunicación pueden construirse con láseres, diodos emisores de luz y fibras de vidrio, examinaremos cómo la información de una fuente, verbigracia un teléfono, una cámara de televisión o un ordenador, se transforma en una señal útil para su transmisión óptica. En los sistemas convencionales de transmisión "analógicos", la onda de la señal original se utiliza para modular la amplitud de la energía entrante en la línea de transmisión, en este caso, la amplitud del rayo de luz emergente de una fuente de luz y entrante en una fibra de vidrio. Del otro extremo de la fibra la luz sale a un fotodetector, que convierte las variaciones de la intensidad de la luz en la señal eléctrica correspondiente. Se amplifica la señal cuanto se necesite para reproducir la forma de onda eléctrica de entrada, para su presentación al oído, ojo o un dispositivo inanimado como un ordenador.

Incluso en las mejores fibras se pierde algo de luz por absorción y esparcimiento, de suerte que la energía de la señal luminosa decrece geométricamente a medida que la señal se propaga desde la fuente hasta el detector. Por ejemplo, si la energía de la señal luminosa cae a la mitad de su valor original en un kilóme-



correspondiente señal eléctrica, que se amplifica luego para reproducir la forma de onda original. En la propagación a través de la fibra óptica, la intensidad de la señal decrece geométricamente con la distancia. La señal comienza a degradarse y distorsionarse, con lo cual la forma de onda reconstruida no es igual a la original. Una manera eficaz de combatir tal degradación consiste en utilizar sistemas de codificación digital.



típico, a la amplitud de la onda, en cada punto de muestreo, se le asigna un valor entre 0 y 255, lo cual requiere una secuencia de ocho dígitos binarios, o bits (ya que 256 es igual a 2^8). Por tanto, para muestrear una onda vocal de un segundo de duración, el sistema digital utilizará 64.000 bits (8000 muestras por ocho bits cada muestra). En los sistemas de comunicación ópticos de Chicago, se diseñan las fuentes de luz utilizadas como transmisores para poder generar 44,7 millones de impulsos por segundo, de suerte que están capacitadas para transmitir más de 650 señales de audio simultáneamente. Aunque los impulsos de luz se degraden en su paso a través de la fibra, los impulsos pueden regenerarse fácilmente (un impulso es presencia o ausencia), así como utilizarse para reconstruir la forma de onda original con alta fidelidad.

tro de recorrido, la energía caería a una cuarta parte del valor original al final del segundo kilómetro, y así sucesivamente. Por tanto, para transmisiones a largas distancias, la potencia de fuente de luz debería ser máxima, y máxima también la sensibilidad del detector.

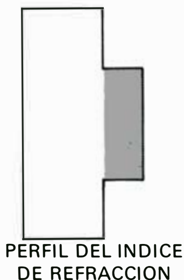
Los láseres de alta intensidad y los fotodetectores ultrasensibles del tipo "avalancha", esto es, detectores en que cada fotón entrante origina una avalancha de electrones, son los que mejor se adaptan actualmente a estos requisitos. Hay que notar, sin embargo, que el alcance de transmisión máxima depende mucho más de las pérdidas en las fibras que de la potencia de la fuente o de la sensibilidad del detector. Por ejemplo, disminuyendo las pérdidas por un factor de dos se dobla el alcance, mientras que incrementando la potencia de la fuente por un factor de dos normalmente se incrementará el alcance en un 10 por ciento (con más precisión, para la misma longitud de la fibra, las pérdidas se incrementarían por un factor de dos).

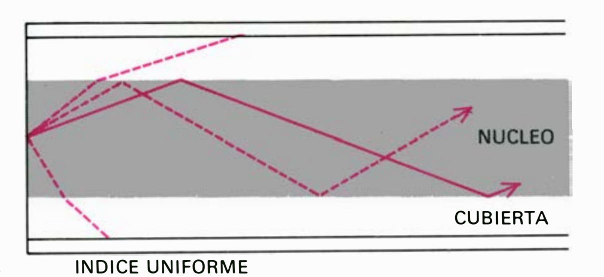
La principal desventaja del sistema de transmisión analógico descrito estriba en que la señal modulada en su amplitud se distorsiona de alguna manera durante el paso a través de la fibra, y cierta cantidad de distorsión resulta inevitable, con la secuencia de que la distorsión estará superpuesta a la señal extraída y amplificada en el receptor.

Según se sabe, uno de los medios más eficaces para transmitir una señal libre de distorsión es codificar la señal en forma digital antes de transmitirla. Lo cual se lleva a cabo muestreando la amplitud, o altura, de la señal electrónicamente, a intervalos regulares. Si se desea representar la onda exactamente, debe muestrearse a una velocidad doble de la de su componente de frecuencia máxima. Por consiguiente, una señal de audio con una frecuencia máxima de 4000 hertzios se representará bien si se muestrea 8000 veces por segundo. Las mediciones de las distintas muestras se codifican en forma binaria, representada por una serie de 1 y 0. Los números binarios se transmiten entonces de acuerdo con un código preestablecido. Por ejemplo, 1 puede transmitirse como un impulso de luz y 0 por su ausencia. En el receptor, los impulsos se detectan y utilizan para reconstruir la onda original.

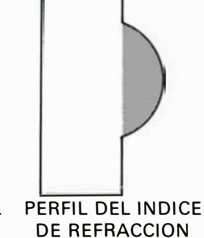
La ventaja más importante de la transmisión digital se pone de manifiesto cuando se trabaja con señales débiles. Cada detector tiene un ruido interno inherente que deforma la señal entrante en el detector, en mayor o menor grado; de ahí que los ingenieros de comunicaciones hablen de la relación señal-ruido. Las relaciones se miden en escala logarítmica de base 10 y en unidades de decibelio. Un decibelio se define como 10 veces el logaritmo de la razón de dos niveles de potencia. Por ejemplo, una relación señal-ruido de 20 decibelios significa que el nivel de la señal es 100 veces más alto que el nivel de ruido. Ya que











LAS FIBRAS OPTICAS se construyen de suerte que el núcleo tenga un índice de refracción mayor que la cubierta. La mayoría de los rayos de luz que chocan contra la superficie de separación entre las dos regiones quedan pues reflejados hacia el núcleo. A menos que la fibra tenga curvas agudas, los rayos van de lado a lado indefinidamente. Si el índice de refracción del núcleo es uniforme *(izquierda)* los rayos que se reflejan suelen seguir un camino más largo, y llegan más tarde que los rayos que sufren pocas reflexiones. Este defecto, llamado dispersión

los impulsos digitales son presencia o ausencia, pueden detectarse con una probabilidad de error baja, incluso con acompañamiento de ruido significativo. Por ejemplo, con una relación señal-ruido de 21 decibelios, solamente un impulso de entre mil millones se perderá en el ruido de fondo. Para señales analógicas, cualquier ruido tiende a distorsionar el mensaje; por consiguiente, si se desea reproducir satisfactoriamente la señal, la relación señal-ruido debe ser muy superior a los 21 decibelios. Generalmente se necesita una relación señal-ruido de 60 decibelios, esto es, una señal que resulte ser un millón de veces mayor que el ruido.

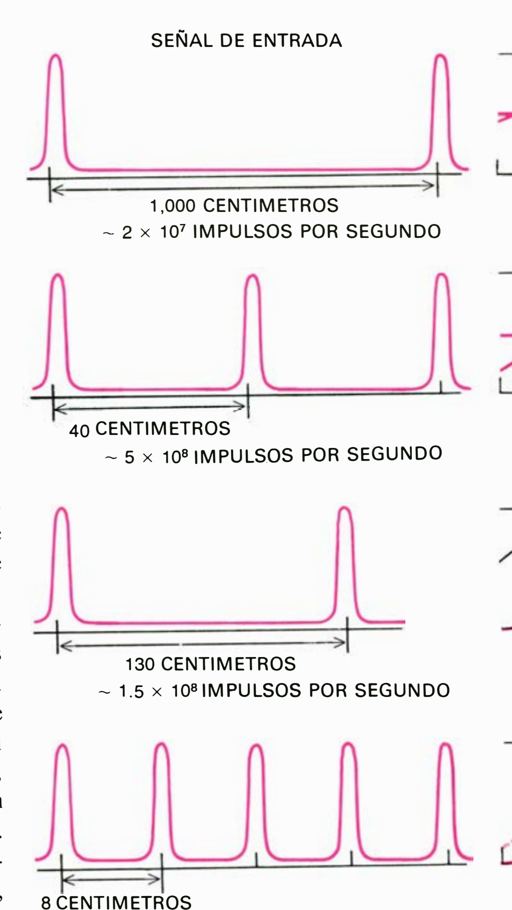
Las tolerancias de ruido más altas de los sistemas de transmisión digital suponen que las señales digitales pueden transmitirse a mayores distancias que las señales analógicas, antes de que sea necesaria la amplificación. Otra ventaja de la transmisión digital se apoya en la facilidad con que los impulsos digitales pueden detectarse y regenerarse. Puesto que las distorsiones menores de la muestra son irrelevantes, los impulsos debilitados pueden detectarse y regenerarse sin imponer requisitos rigurosos al amplificador.

Cada vez es más frecuente en telefonía el envío de las señales de audio en sistemas de transmisión por cable o microondas mediante impulsos digitales. La señal de audio es muestreada 8000 veces por segundo, representándose con ocho dígitos binarios de altura de cada muestra. Ya que los ocho dígitos binarios pueden representar 2^8, esto es 256, niveles de amplitud, proporcionan una especificación exacta de la muestra de la onda. Lo cual significa que, para reproducir la voz original, que tiene un ancho de banda de frecuencia de 4000 hertzios, el sistema digital debe estar preparado para transmitir 64.000 impulsos por segundo. El gran ancho de banda de los sistemas ópticos permite la utilización de anchos de banda mayores a costa de reducir la relación señal-ruido, normalmente alta, para aumentar el alcance del sistema, es decir, la distancia a la cual la señal puede llegar sin ser regenerada.

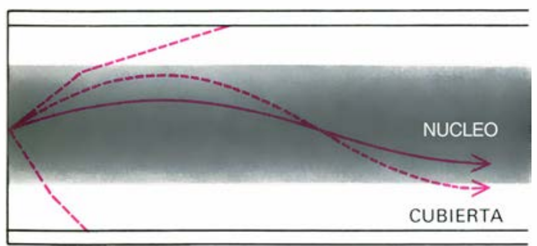
Vemos, pues, que en sistemas de comunicación óptica el alcance depende de la potencia de la fuente, de la atenuación por unidad de longitud de la fibra, del nivel de ruido del detector y del tipo de modulación o codificación que se utilice. La capacidad del sistema en ancho de banda, impulsos por segundo o cualquier otra medida de la capacidad de información, depende de la velocidad con que la fuente pasa de estar encendida a estar apagada, de la velocidad de respuesta del detector y también de las características de ensanchamiento del impulso.

Suelen emplearse dos clases de fuente de luz. La primera es una versión refinada de los LED, que se usa en las unidades de presentación visual de las calculadoras de bolsillo. Para comunicaciones ópticas, lo que se necesita es una fuente que no sólo sea mucho más intensa que las de presentación visual, sino comparable también en tamaño de la fibra óptica, que tiene por diámetro unas pocas centésimas de milímetro. Al diseñar el LED para comunicaciones ópticas, se prevé un pequeño rebaje en la cara del LED para aproximar, lo máximo posible, la fibra a la región activa de la unión semiconductora, donde se origina la luz. Como las fibras presentan menos pérdidas en la zona infrarroja del espectro, se selecciona un material semiconductor que emita radiación infrarroja. Los LED hechos de arseniuro de galio, que emiten a una longitud de onda de 0,8 micras, dan buen resultado, pero lo deseable es que la longitud de onda de emisión fuese algo más larga. Se está investigando en materiales semiconductores que prometen una longitud de onda mejor que las fibras ópticas.

El otro tipo de fuente de luz es el láser de diodo semiconductor, que tiene una estructura más compleja que el diodo emisor de luz. No mayor que un grano de arena, un diodo láser consta de varias capas de material semiconductor, cada una de composición diferente. La estructura en bocadillo ayuda a establecer las condiciones necesarias para la acción



DOS TIPOS DE DISPERSION, o ensanchamiento de impulso, son los que hay que considerar en las comunicaciones ópticas. El primer tipo, dispersión modal, aparece en la parte superior. El conjunto de diagramas muestra cómo el ensanchamiento del impulso limita la velocidad del mismo y la capacidad de las fibras ópticas. En cada caso, se selecciona la velocidad del impulso para que el ensanchamiento alcance la mitad del intervalo entre impulsos. Para una fibra con un núcleo de índice de refracción uniforme *(a)*, la diferencia entre el tiempo de llegada del flanco inicial y el de la llegada del



modal, se corrige fabricando una fibra cuyo índice de refracción aumente hacia el eje ***(derecha).*** **Los rayos que se desvíen del eje viajarán más rápidamente que los que se separen menos.**

láser; proporciona una región en donde se confinan los portadores que emiten luz cuando se recombinan; ayuda, asimismo, a guiar la luz en una dirección determinada.

Fue difícil obtener sucesivas capas de material en un láser sin dañar su estructura cristalina. Los primeros dispositivos se significaron por la rápida degeneración de la eficiencia de luz emitida; algunos fallaron a las pocas horas. Se fueron desarrollando nuevas técnicas gradualmente y ya puede fabricarse la estructura compuesta sin introducir imperfecciones en las capas cristalinas. Las pruebas de envejecimiento acelerado demuestran que los dispositivos desarrollados más recientemente resistirán durante varios años a temperaturas ambientales. Se espera que, en un futuro próximo, los diodos láser resulten tan fiables como otros dispositivos de estado sólido.

La fuente láser tiene dos ventajas principales. La primera es su direccionalidad. A causa de la emisión estimulada del láser se obtiene un rayo estrecho que hace posible acoplar directamente una gran fracción de radiación al extremo de una fibra óptica. La segunda ventaja es su pequeño margen cromático de longitud de onda, que es típico de una fuente láser. En la propagación a través de una fibra óptica, los rayos de longitudes de onda diferentes se propagan a velocidades ligeramente distintas; por ello, el ensanchamiento de los impulsos en una fibra óptica varía directamente con el ancho de banda de las longitudes de onda transmitidas. Las fuentes láser están capacitadas para transmitir impulsos, sobre una distancia dada, a velocidades más altas que los LED, que emiten una banda de longitudes de onda más ancha. El espectro de un diodo láser típico tiene un ancho de 20 angstroms, comparado con los 350 angstroms de un diodo emisor de luz del tipo utilizable para comunicaciones por fibra óptica. Después de propagarse un kilómetro a través de una fibra óptica, el impulso láser mostrará una dispersión en tiempo de 200×10^{-12}

CARACTERISTICA DE DISPERSION

SEÑAL DE SALIDA

~ 500 CENTIMETROS POR KILOMETRO

500 CENTIMETROS

a. FIBRA DE INDICE UNIFORME

~ 20 CENTIMETROS POR KILOMETRO

20 CENTIMETROS

b. FIBRA DE INDICE GRADUAL

~ 65 CENTIMETROS POR KILOMETRO

65 CENTIMETROS

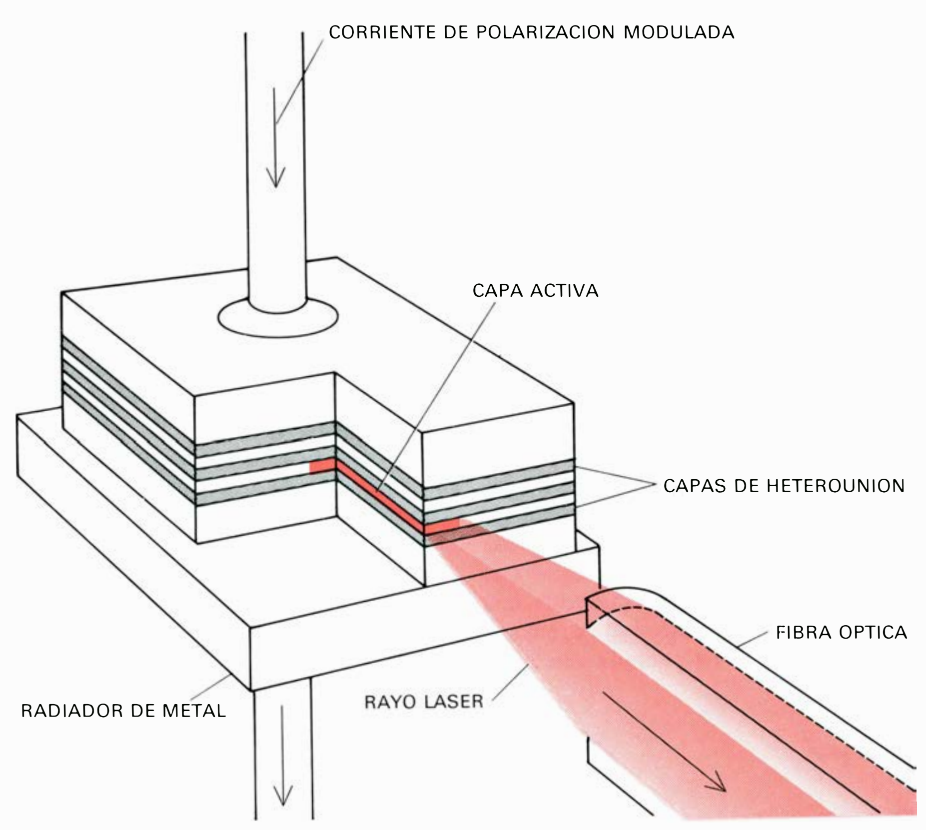
c. FUENTE DE LUZ LED (ANCHO DE BANDA: 350 ANGSTROMS)

~ 4 CENTIMETROS POR KILOMETRO

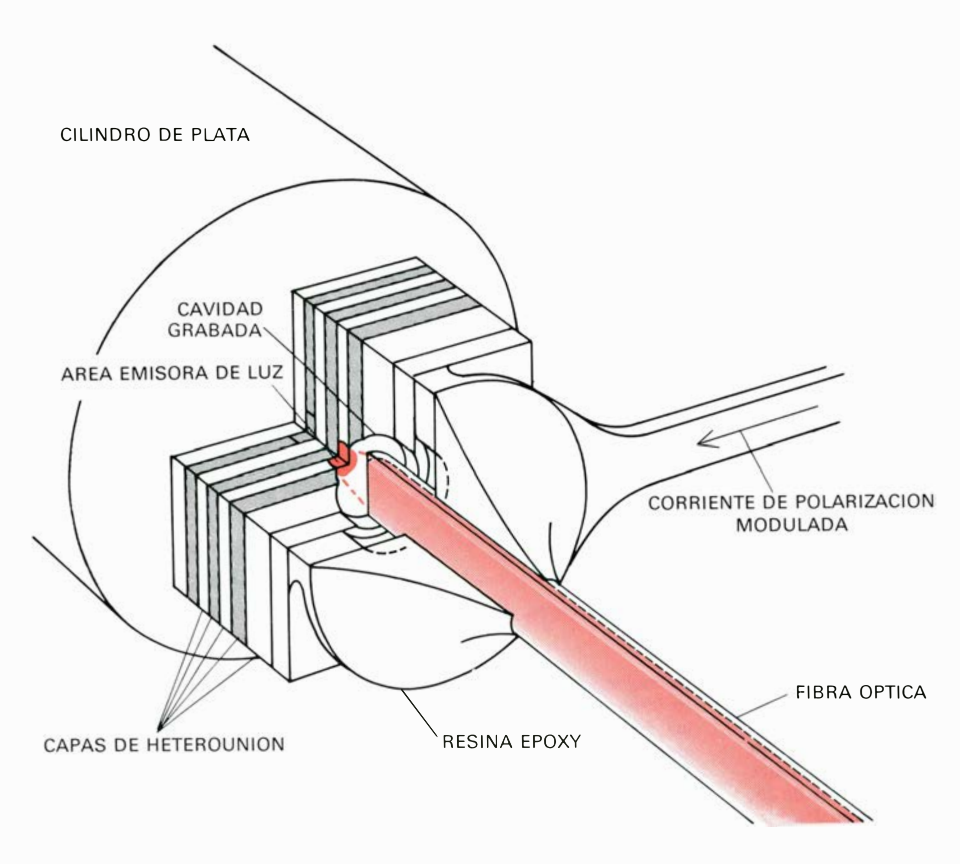
4 CENTIMETROS

d. FUENTE DE LUZ LASER (ANCHO DE BANDA: 20 ANGSTROMS)

final de un impulso, arbitrariamente estrecho, puede tomar el valor 25×10^{-9} segundos por kilómetro, que equivale a 500 centímetros para la velocidad de propagación de la luz en el vidrio. A fin de que el ensanchamiento se mantenga a la mitad del intervalo, la velocidad del impulso no puede exceder de 2×10^{7} impulsos por segundo. Las fibras con índice de refracción gradual reducen la dispersión modal en un factor de 25, en este caso a 10^{-9} segundos por kilómetro, o sea a 20 centímetros ***(b).*** **De ahí que puedan transmitirse impulsos arbitrariamente agudos a una velocidad de 5×10^{8} impulsos por segundo. (Las distancias de ensanchamiento se trazan en unidades logarítmicas; los impulsos en la figura de la izquierda están separados por una distancia doble de la de ensanchamiento.) El segundo tipo de ensanchamiento de impulso, dispersión de longitud de onda** ***(c y d),*** **se debe a que la velocidad de las ondas electromagnéticas varía con el índice de refracción del medio: la más alta en frecuencia se propaga a velocidad más baja. Los diodos emisores de luz de alta intensidad (LED) que están siendo utilizados en comunicaciones ópticas tienen un ancho espectral de 350 angstroms, centrados en la zona infrarroja del espectro a una longitud de onda de 0,82 micras. Su ancho de banda es aproximadamente igual al intervalo entre verde y amarillo del espectro visible. En una fibra de índice gradual con dispersión modal cero, un simple impulso de un LED se ensancharía casi unos 65 centímetros por kilómetro, limitando así la velocidad de la señal a $1{,}5 \times 10^{8}$ impulsos por segundo** ***(c).*** **Con una fuente de luz láser** ***(d),*** **dotada de un ancho de banda mucho más estrecho (20 angstroms), la dispersión de longitud de onda alcanzará sólo cuatro centímetros, permitiendo una velocidad de señal igual a 3×10^{9} impulsos por segundo.**



LA FUENTE DE LUZ LASER utilizada en comunicaciones ópticas es del tamaño de un grano de arena. La acción láser (amplificación de luz por emisión estimulada de radiación) tiene lugar dentro de una estructura de "heterounión", formada por capas semiconductoras de arseniuro de galio y de arseniuro de galio-aluminio con características electrónicas rigurosamente controladas. El rayo láser emerge de una de las capas de arseniuro de galio y se propaga cuarenta micras a través de un medio gaseoso, que se encuentre herméticamente cerrado, antes de penetrar en la fibra óptica.



LOS DIODOS EMISORES DE LUZ, constituidos asimismo por capas de heterounión, son más simples, baratos y más fiables que los láseres. Sirven cuando no se necesitan anchos de banda estrechos y en distancias de transmisión donde basta su potencia media (0,1 milivatios).

segundos, equivalente a una dispersión en distancia de cuatro centímetros a la velocidad con la que la luz se propaga a través del vidrio. Para una fuente LED la dispersión es aproximadamente 20 veces mayor. Esta dispersión, debida a una carencia de zonas de pureza espectral, limita de forma notable la velocidad de los impulsos y recorta, por tanto, la capacidad de información del sistema de comunicación óptica. Otra limitación importante consiste en un tipo de ensanchamiento de impulso debido a dispersión modal; ésta obedece a que ciertos rayos que entraron en la fibra óptica viajan por caminos ligeramente más largos que otros rayos. Como se verá, la dispersión modal puede reducirse sustancialmente, aunque no eliminarse por completo.

Para obtener la transparencia extraordinaria que se necesita para guiar la luz, las fibras ópticas se diseñan de suerte que la luz nunca se propague cerca de la superficie externa de la fibra, donde el polvo, rayados o contacto con otras superficies causarían pérdidas importantes. Cada fibra consta de tres capas. La capa externa es una envoltura de plástico, que protege de rayados y desgastes, que podrían debilitar la fibra y provocar una ruptura por tensión. Dentro de la capa protectora, la fibra de vidrio tiene un núcleo con un índice de refracción ligeramente más alto que el de la cubierta que le rodea. Debido a que el índice de refracción es más alto, los rayos que entran por el extremo de la fibra formando un ángulo pequeño con el eje central son reflejados hacia el núcleo cuando chocan contra la superficie de separación entre el núcleo y la cubierta. Los rayos que entran en la fibra formando ángulos grandes respecto al eje central escaparán simplemente, sin ser reflejados. De las consideraciones geométricas puede deducirse que si un rayo se refleja hacia el núcleo en su primer choque con la superficie de separación, continuará confinado indefinidamente, con tal de que no haya curvas agudas en la fibra. Tales curvas pueden ser evitadas encerrando cuidadosamente varias fibras en un cable de envoltura rígida.

Se puede visualizar ya la fuente de la dispersión modal: un rayo de luz que entre paralelo al eje central de una fibra recorrerá una distancia más corta que otro rayo que entre con un cierto ángulo y se propague a través de la fibra reflejándose en la superficie de separación. En consecuencia, un impulso de luz formado por una combinación de rayos aumentará su duración.

Para vencer el alargamiento del impulso, muchas fibras del mercado tienen un núcleo cuyos índices de refracción están graduados, o modelados para compensar las diferentes distancias de recorrido de los rayos. En tales fibras, el índice de refracción decrece con la distancia radial al centro de la fibra. En las regiones de índice de refracción más bajo, la luz se propaga con mayor rapidez. Al hacer que el índice de refracción disminuya radialmente, todos los rayos pueden llegar a su destino más o menos al mismo tiempo. En una fibra de índice de refracción uniforme, el alargamiento del impulso es aproximadamente 25×10^{-9} segundos por kilómetro, que equivale a unos 500 centímetros. Las fibras de índice gradual, sometidas ahora a un proceso de experimentación, reducen esta dispersión en un factor de 25, y en muestras de laboratorio se han obtenido mejoras con un factor de 100.

Las primeras fibras de alta transparencia fueron fabricadas por la Corning Glass Works con un material cuyo componente principal era el dióxido de silicio. La primera fibra de índice de refracción gradual fue fabricada con éxito por la Nippon Sheet Glass Co., Ltd. En el proceso desarrollado por los Bell Laboratories, se obtiene una fibra de índice gradual calentando y colapsando un tubo de 0,9 metros de cristal de cuarzo, cuyo interior se ha recubierto previamente con docenas de capas de dióxido de silicio dopado con germanio, controladas rigurosamente. Cada capa tiene un espesor aproximado de una centésima de milímetro. El tubo se colapsa, y así se forma un cilindro sólido llamado preforma; posteriormente, éste se convierte, por estirado, en una fibra de varios kilómetros de longitud.

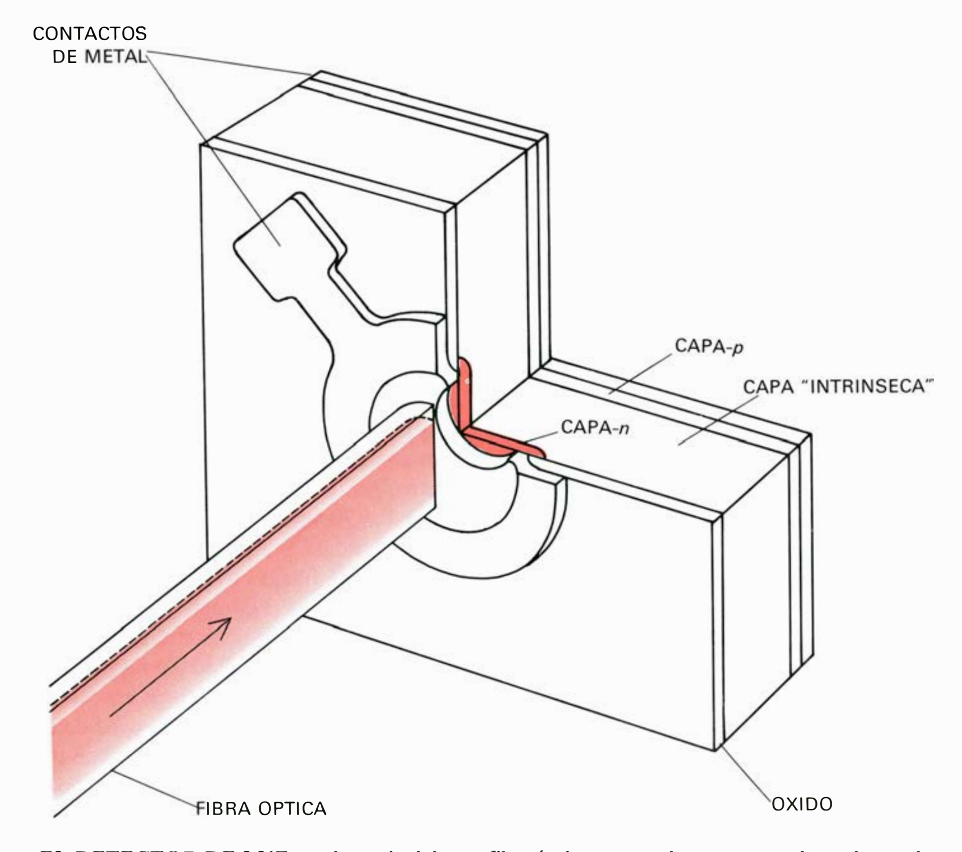
En las mejores muestras de fibras las pérdidas de transmisión pueden alcanzar un decibelio por kilómetro, que equivale a la transmisión de un 80 por ciento de la energía de entrada. Sin embargo, para la frecuencia de trabajo de las fuentes de luz disponibles, no pueden rebajarse hasta ese punto las pérdidas. En los procesos de fabricación normales, la media real de las pérdidas suele hallarse por los cuatro o cinco decibelios por kilómetro, que supone una transmisión del 30 por ciento de la energía de entrada. Con estos valores, no obstante, los impulsos de luz láser pueden transmitirse a 14 kilómetros de distancia, antes de que sea necesaria una amplificación. (A esta distancia, sólo se conserva un 10^{-7} de la energía de entrada.) A medida que las fuentes y los detectores "sintonicen" con la región del espectro de longitudes de onda algo superiores a la micra, donde las fibras ofrecen pérdidas mínimas, y a medida que se mejore la calidad de las fibras, la distancia entre amplificadores podrá prolongarse más allá de los 14 kilómetros.

Las fibras ópticas del grosor de un cabello se cablean fácilmente. Una vez revestidas como protección contra la humedad, desgaste y pérdidas debidas a curvaturas, las fibras se ensamblan horizontalmente, formando cintas de 12 fibras cada una, que se codifican mediante colores. Hasta una docena de cintas se agrupan en un cable, que protege y defiende a las fibras contra daños mientras está en servicio. Se derrochó considerable ingenio para inventar métodos de empalme eficientes. Por último se desarrolló una técnica que permite alinear los extremos de las fibras de un cable con una precisión de dos micras.

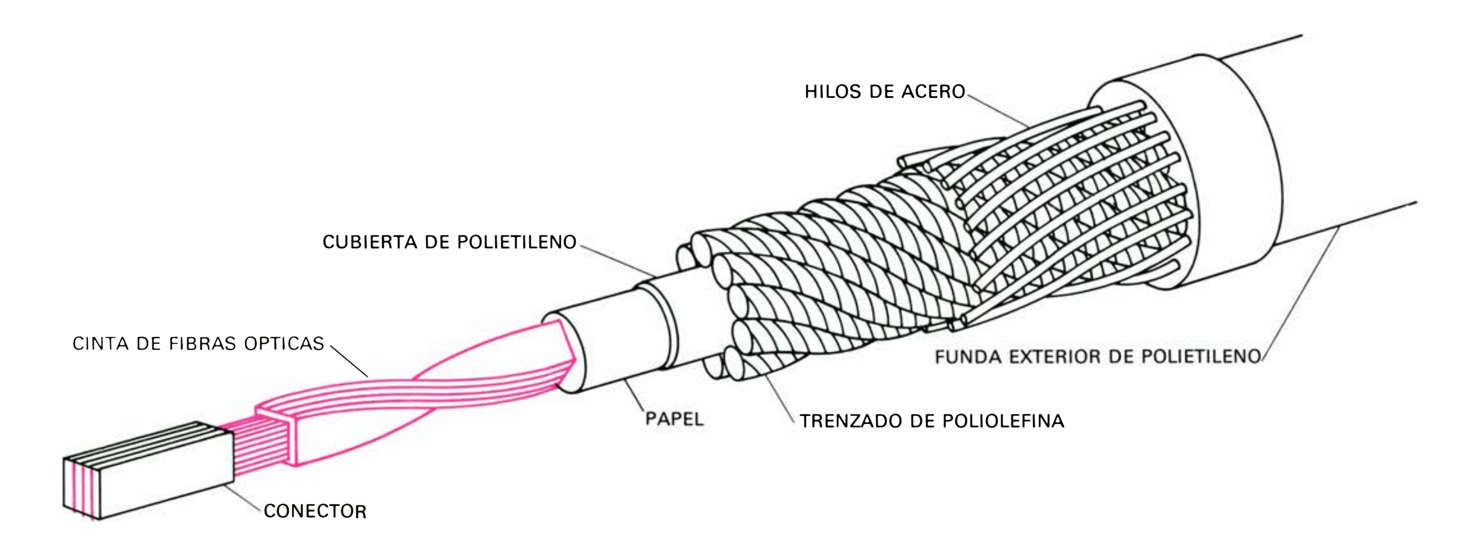
Las fibras ópticas ofrecen muchas ventajas sobre la transmisión por conductores metálicos. Puesto que la luz en un sistema de transmisión de fibra óptica está estrechamente confinada al núcleo interno de cada fibra, las señales no pueden dispersarse hacia las fibras adyacentes para dar lugar a la diafonía. Además, al no estar afectadas las fibras ópticas por interferencias eléctricas de otras fuentes, los sistemas ópticos mostrarían ventajas en el transporte de información en centros eléctricamente ruidosos, por ejemplo, entre aparatos de conmutación en centrales telefónicas.

Los cables para comunicaciones ópticas ofrecen diferentes ahorros en materiales comparados con los cables metálicos de capacidad equivalente. En la actualidad, las fibras ópticas cuestan bastante más que el hilo de cobre, pero es lo normal cuando un nuevo producto, tecnológicamente complejo, entra por primera vez en un sistema de producción.

Así como hay dos clases de fuentes para comunicaciones ópticas, se utilizan también dos tipos de detectores. Son dispositivos de estado sólido. Uno es un dispositivo simple del tipo unión, cono-



EL DETECTOR DE LUZ en el terminal de una fibra óptica genera electrones cuando es alcanzado por fotones. El detector más simple es el fotodiodo PIN, que aparece en esta ilustración. El término PIN significa que el detector consta de una región "intrínseca" (*i*) de silicio ligeramente tipo *p*, situada entre capas de silicio tipo *p* y tipo *n*. (Un material tipo *p*, o tipo positivo, es deficiente en electrones; un material tipo *n*, o tipo negativo, tiene un exceso de electrones.) Los fotones de luz que son absorbidos en la región *i* del fotodiodo dan lugar a electrones y huecos (ausencia de electrón) que se mueven bajo la acción de un campo eléctrico uniforme para producir una corriente eléctrica. Un fotodetector más complejo, conocido como fotodiodo de avalancha, incorpora una capa adicional de material tipo *n* que proporciona un proceso de amplificación para aumentar la señal eléctrica. Los fotodiodos tienen un ruido propio que se incrementa con la velocidad de trabajo. Cuando un detector PIN recibe señales a una velocidad de 10^8 impulsos por segundo, su ruido propio es de 10^{-9} vatios. Para el detector de avalancha el ruido propio es inferior en un factor de 10.



EL CABLE DE FIBRA OPTICA del tipo utilizado en la demostración del año pasado en Atlanta, Georgia, contiene 144 fibras de vidrio agrupadas en 12 cintas de 12 fibras cada una. Las cintas, apiladas una junto a otra, están recubiertas por varias capas de material protector; todas están contenidas en una cubierta de polietileno con hilos de acero. Operando a una velocidad de impulso de 44,7 millones de ciclos por segundo, un par de fibras pueden transportar 672 conversaciones telefónicas bidireccionales, una conversación de vídeo bidireccional o un equivalente de información de otro tipo. Las secciones de cable están equipadas con conectores prefabricados, que se cierran con broches de precisión.

cido como detector PIN, parecido a una célula solar, en el cual los fotones de luz generan una corriente eléctrica. (Las letras P, I y N representan las propiedades electrónicas de los semiconductores empleados en la unión del detector.) El otro dispositivo es el fotodetector de avalancha. Todos los detectores de señal tienen un ruido de fondo que incrementa en proporción a su velocidad de operación. Por ejemplo, el ruido de fondo en un detector PIN sube desde 10^{-11} vatios, cuando el dispositivo opera a un megabit por segundo, a 10^{-9} vatios a 100 megabits por segundo. A las mismas velocidades de operación, el ruido de fondo en un detector de avalancha es inferior en un factor de 10. Ello implica que las distancias de transmisión para sistemas de baja velocidad son mayores que en sistemas de alta velocidad. En comunicaciones ópticas, el detector de señal es la primera etapa de un módulo receptor que contiene la circuitería necesaria para adaptar las señales para su transmisión a través de la red de telecomunicación existente.

Séanos permitido considerar en conjunto las informaciones sobre las propiedades de fuentes, los detectores y las fibras para determinar qué tipos de comunicación resultan idóneos. En primer lugar hay que calcular el alcance de un sistema de baja velocidad, capaz de transmitir 10^6 bits por segundo. Para que la detección se efectúe con pocos errores, la señal de llegada al detector debe ser 100 veces mayor que el ruido interno de éste. Si se emplea un fotodetector de avalancha, la señal de llegada debe tener un nivel de potencia de 10^{-10} vatios, por lo menos. Para el alcance máximo podríamos elegir un láser con una potencia de salida de 10^{-3} vatios, con preferencia a un LED, que es un orden de magnitud menos potente. Como hemos visto, con codificación digital la atenuación máxima permitida para la luz que pasa a través de la fibra es de un factor de 10^7, es decir, 70 decibelios. Puesto que las fibras disponibles tienen una atenuación inferior a los cinco decibelios por kilómetro, podemos esperar una transmisión óptima para una distancia de 14 kilómetros, antes de recurrir a la amplificación. (Si se consiguieran fibras con atenuación de un decibelio por kilómetro, el alcance podría prolongarse hasta los 70 kilómetros.) En la práctica, se duda de la viabilidad de los tramos continuos de fibras cuya longitud supere a unos cuantos kilómetros. Así pues, la pérdida adicional introducida en la unión de dos fibras debe sumarse a las pérdidas. Hoy, los conectores especiales de tipo clavija introducen una pérdida de 0,5 decibelios. Si se necesitaran seis conectores para una ruta de 14 kilómetros, la pérdida adicional llegaría a tres decibelios. (La pérdida total podría mantenerse en 70 decibelios acortando la ruta en tres quintos de un kilómetro.)

Mediante la adecuada selección de una fuente, un detector y una fibra ¿cual sería la capacidad de información del sistema óptico? Se trata, obviamente, de transmitir a la velocidad de bit más alta posible; para ello tienen que entrar en consideración varios factores. Como hemos visto, el nivel de ruido del detector aumenta con la velocidad de bit. Así, si la potencia de la señal fuese justamente la adecuada para la transmisión a 10^6 impulsos por segundo, habría de incrementarse en un factor de 100 para la transmisión a 10^8 bits por segundo. Además, como los impulsos obtenidos son más cortos, y más próximos sus ensanchamientos, a medida que se propagan por la fibra, ésta se convierte en un factor de limitación importante.

Para simplificar el cálculo se decide arbitrariamente que el ensanchamiento del impulso no exceda la mitad del intervalo entre dos impulsos sucesivos. En una fibra de índice gradual, el ensanchamiento de impulso debido a la dispersión modal (diferencias en longitudes de camino) alcanza cerca de 10^{-9} segundos por kilómetro, lo cual significa que si se intenta transmitir 10^9 impulsos por segundo, el ensanchamiento equivale a un intervalo entero entre picos de impulsos. Por tanto, para mantener una separación de medio intervalo, la velocidad de la señal no puede exceder de $0{,}5 \times 10^9$ (o lo que es lo mismo 5×10^8) impulsos por segundo. Este es el límite si tenemos una fuente láser, que es casi monocromática, con lo cual el ensanchamiento del impulso debido a la dispersión de la longitud de onda puede ser despreciado.

Ahora bien, si elegimos un LED como fuente de luz, la dispersión de longitud de onda se convierte en el factor que limita la velocidad de la señal. Para un LED, la dispersión de longitud de onda se aproxima a $3{,}5 \times 10^{-9}$ segundos por

kilómetros, 3,5 veces mayor que la dispersión modal. Para mantener el ensanchamiento por debajo de medio intervalo entre impulsos sucesivos, la velocidad de la señal, con una fuente LED, debe ser ligeramente inferior a un tercio de la velocidad permisible con un láser, exactamente $1,4 \times 10^8$ impulsos por segundo. Conforme la distancia de transmisión vaya aumentando, debe ir reduciéndose, proporcionalmente, la velocidad. Por ejemplo, para una distancia de transmisión nominal de 10 kilómetros, la velocidad para la fuente láser habrá de reducirse diez veces, a 5×10^7 impulsos por segundo, o aproximadamente la velocidad elegida para la instalación de Chicago ($4,47 \times 10^7$). Estos cálculos simples ilustran qué es lo que la tecnología actual permite realizar, y ofrecen asimismo una panorámica de los tipos de diseño que pueden seguirse, habida cuenta del alcance, la capacidad y la complejidad de los dispositivos. Indudablemente, se obtendrán mejoras significativas en el futuro.

Hay muchas áreas donde la aplicación de la nueva tecnología de luz guiada se hace prometedora. Por ejemplo, las señales de televisión pueden transportarse fácilmente a través de una fibra única, presentando nuevas ventajas para la aplicación de este servicio al ocio y a los negocios. Fibras empotradas, en cables, pueden proporcionar a los edificios un buen sistema de comunicación interna. Asimismo, distintas partes de los ordenadores pueden interconectarse mediante fibras. Es en telefonía, sin embargo, en donde se espera ocurran las primeras aplicaciones de importancia, que, además, sirvan de modelo.

La mayoría del cableado de cobre que interconecta las centrales telefónicas urbanas de nuestros días está tendido en canalizaciones subterráneas, cuyo espacio está muy solicitado. Añadir nuevas canalizaciones resulta costoso e inconveniente. Los sistemas de comunicaciones ópticas, con sus altas capacidades y pequeño tamaño, pueden hacer mejor uso de las canalizaciones ya existentes y ayudar a aplazar la necesidad de otras nuevas. Más aún, ya que las centrales telefónicas adyacentes se encuentran a menos de siete kilómetros, los sistemas ópticos no requerirán amplificadores en las cajas de registro para regenerar las señales a lo largo de una ruta típica.

Antes de terminarse la instalación de Chicago, los Bell Laboratories y la Western Electric Company probaron el año pasado en Atlanta un sistema óptico prototipo, que simulaba las condiciones de campo. Se colocaron dos cables de fibra óptica de 640 metros de longitud, cada uno de los cuales contenía 144 fibras, a través de las canalizaciones normalizadas; se sometieron a experimentación simulando un enlace de telecomunicación urbano típico. El trabajo de instalación no rompió ninguna de las fibras ópticas; la colocación, que requería la formación de curvas agudas, tampoco degradó el comportamiento de las fibras ópticas. En el sistema de Chicago, cada par de fibras transporta el equivalente a 672 canales vocales bidireccionales. Las fuentes de luz eran láseres de arseniuro de galio-aluminio, que operaban a una velocidad de 44,7 millones de bits por segundo. En el extremo receptor, los impulsos de luz se convertían en señales eléctricas por fotodetectores de avalancha.

Como parte del experimento de Atlanta, los extremos de algunas fibras se unieron para crear una comunicación continua con un recorrido de 70 kilómetros. Con la ayuda de 11 regeneradores, o amplificadores, se llevó a cabo una transmisión, virtualmente libre de error, en una dirección a lo largo de todo el camino, durante un período prolongado. La instalación de Chicago sigue estrechamente el sistema experimental de Atlanta, con la salvedad de que emplea LED, en vez de láser, como fuente de luz.

Aparte de algunas alusiones sobre la reducción de la pérdida de fibra en el futuro, todo lo descrito está basado sobre la tecnología en curso. Sería contradecir las experiencias anteriores creer que no se presenciarán desarrollos posteriores. Por ejemplo, investigadores de la industria y de la universidad están experimentando con óptica integrada, que comprende técnicas para procesar señales de luz dentro de películas delgadas (equivalente óptico de los circuitos microeléctronicos integrados). Quizá tales circuitos ópticos eliminen algún día la necesidad de convertir impulsos de luz en señales eléctricas, y viceversa, en los amplificadores colocados a lo largo del camino de transmisión. Además, el trabajo teórico y experimental se orienta sobre la posibilidad de conmutación directa de impulsos de luz, evitando la necesidad de convertir estas señales en su equivalente eléctrico. La esperanza está puesta en el desarrollo de conmutadores ópticos que reemplacen a los dispositivos electromecánicos y electrónicos actuales, de tal forma que haga posible conectar llamadas telefónicas en número mayor y a velocidades más altas.

Ciencia y sociedad

El principio de complementariedad

En septiembre de 1927, con ocasión del centenario de la muerte de Alessandro Volta, se celebró en Como, su ciudad natal, un Congreso Internacional de Física. En él participaron unos sesenta físicos procedentes de catorce naciones, entre ellos una docena de premios Nobel y media de futuros premios Nobel. (Participantes españoles fueron el Prof. E. Alcobé, en representación de la Academia de Ciencias de Barcelona, y el Prof. B. Cabrera, en representación de la de Madrid.) Por primera vez en la historia y como signo del progreso eléctrico realizado desde los tiempos de Volta, las sesiones del congreso fueron retransmitidas por radio. Además de las sesiones conmemorativas, que inauguraron y clausuraron el congreso, se dedicaron cinco jornadas a sesiones estrictamente científicas, sobre la electricidad y sus aplicaciones, y también sobre la estructura de la materia, tanto desde el punto de vista experimental como teórico. Como conclusión de la última de esas jornadas, la del 16 de septiembre, Niels Bohr presentó una comunicación célebre, que introdujo en la nueva mecánica cuántica el llamado "principio de complementariedad", y constituyó el segundo episodio de la revolución epistemológica conocida como "Interpretación de Copenhague". (Sobre el primer episodio véase INVESTIGACION Y CIENCIA, julio).

Niels Bohr habló en inglés, y tituló su comunicación "El postulado cuántico y el desarrollo reciente de la teoría atómica". Tras una breve introducción en que recuerda a "los antiguos maestros que nos han preparado el terreno y nos han proporcionado los instrumentos", Bohr entra directamente en el núcleo de su concepción epistemológica. Traducimos literalmente sus tres primeros párrafos, sacrificando cuanto no creemos esencial a su razonamiento y subrayando lo que juzgamos central en el mismo: "Es característico de la teoría cuántica reconocer una limitación fundamental de nuestras ideas físicas clásicas, al aplicarlas a los fenómenos atómicos... Lo esencial de esta teoría puede expresarse en el llamado *postulado cuántico*, que *atribuye a los procesos atómicos una esencial discontinuidad,* o, mejor, individualidad, totalmente extraña a las teorías clásicas y simbolizada por el cuanto de acción de Planck.

"Este postulado nos obliga a renunciar a una descripción de los procesos atómicos a la vez causal y espacio-temporal. En efecto, nuestra descripción ordinaria de los fenómenos físicos se basa totalmente en la idea de que podemos observar los fenómenos en cuestión, sin perturbarlos apreciablemente... Ahora bien, el postulado cuántico implica que *toda observación de fenómenos atómicos lleva consigo una interacción con el aparato de medida, que no puede despreciarse....*

"Esta situación tiene consecuencias trascendentales. Por una parte, *para definir el estado de un sistema físico,* en su sentido ordinario, es necesario excluir toda perturbación externa. Pero entonces, según el postulado cuántico, *toda observación quedará excluida.* Por otra parte, si *para hacer posible la observación* permitimos ciertas interacciones con aparatos de medida apropiados que no pertenecen al sistema, evidentemente ya no es posible definir el estado del sistema, y *no cabe hablar de causalidad,* en el sentido ordinario de la palabra. La naturaleza misma de la teoría cuántica nos obliga, pues, a considerar *la coordenación espacio-temporal y la exigencia de causalidad,* cuya combinación caracteriza las teorías clásicas, como *dos rasgos complementarios y mutuamente excluyentes de la descripción,* que simbolizan respectivamente la observación ideal y la definición ideal".

Así es como se introduce en el contexto de la nueva física el término "complementario", que Niels Bohr repetirá una decena de veces a lo largo de su comunicación. El sentido literal de ese término es claro: califica rasgos diversos que se armonizan perfectamente en la descripción física clásica, pero que en la nueva física cuántica aparecen como mutuamente excluyentes, y, sin embargo, resultan ambos necesarios para la descripción total del fenómeno. Pero Bohr ha cargado de sentido epistemológico ese término. Su complementariedad es asimétrica, al colocar a niveles distintos los dos rasgos complementarios. Uno corresponde a la "observación" del sistema, que se funda en último término, según había resaltado la física relativista, en la determinación de las coordenadas "espacio-temporales" de un conjunto de sucesos puntuales. El otro corresponde a la "definición del estado" del sistema, que propiamente sólo tiene sentido tratándose de sistemas cerrados, es decir, libres de perturbaciones externas, y que permite predecir la evolución "causal" del sistema (con más propiedad diríamos: su evolución "determinista").

Intentemos concretizar esas ideas tan abstractas. En la descripción física de un sistema clásico, como el sistema solar newtoniano, esos dos rasgos de definición y observación se armonizan perfectamente. Si prescindimos de perturbaciones extraplanetarias, en realidad muy débiles o muy raras, cabe definir el estado actual del sistema solar: posición y velocidad de cada uno de los planetas, cuyas masas suponemos conocidas. Según las leyes deterministas de la mecánica clásica, esto nos permite conocer su evolución ulterior, por ejemplo, predecir eclipses. Esas predicciones tienen valor científico gracias al segundo rasgo, el de la observación, que nos permite, verbigracia, comprobar los eclipses (¡sin perturbar el sistema solar!). Pero esos dos rasgos se excluyen mutuamente cuando se trata de un sistema cuántico, como por ejemplo de un átomo estable, descrito según el modelo atómico que introdujo el propio Niels Bohr en 1913 y modificó sustancialmente la nueva mecánica cuántica de 1925. Definido un estado del sistema, también aquí queda determinada su evolución ulterior. Pero toda observación de su estructura implica una perturbación esencial del sistema, y rompe su evolución determinista. Sin embargo, ambos rasgos de definición y observación han de incluirse, como complementarios, en la descripción física total del sistema atómico.

Este es el sentido primigenio del que vendrá a llamarse "principio de complementariedad" de Bohr, aunque aún no se considere como principio, ni el término "complementariedad" esté todavía definitivamente acuñado. (El propio Bohr en una conferencia ulterior decidirá cambiarlo por el de "reciprocidad", pero pronto desistirá de ese cambio.)

Insinuemos algunos aspectos de la evolución conceptual de este principio, a propósito de otros pasajes de esta comunicación de Niels Bohr al Congreso de Como.

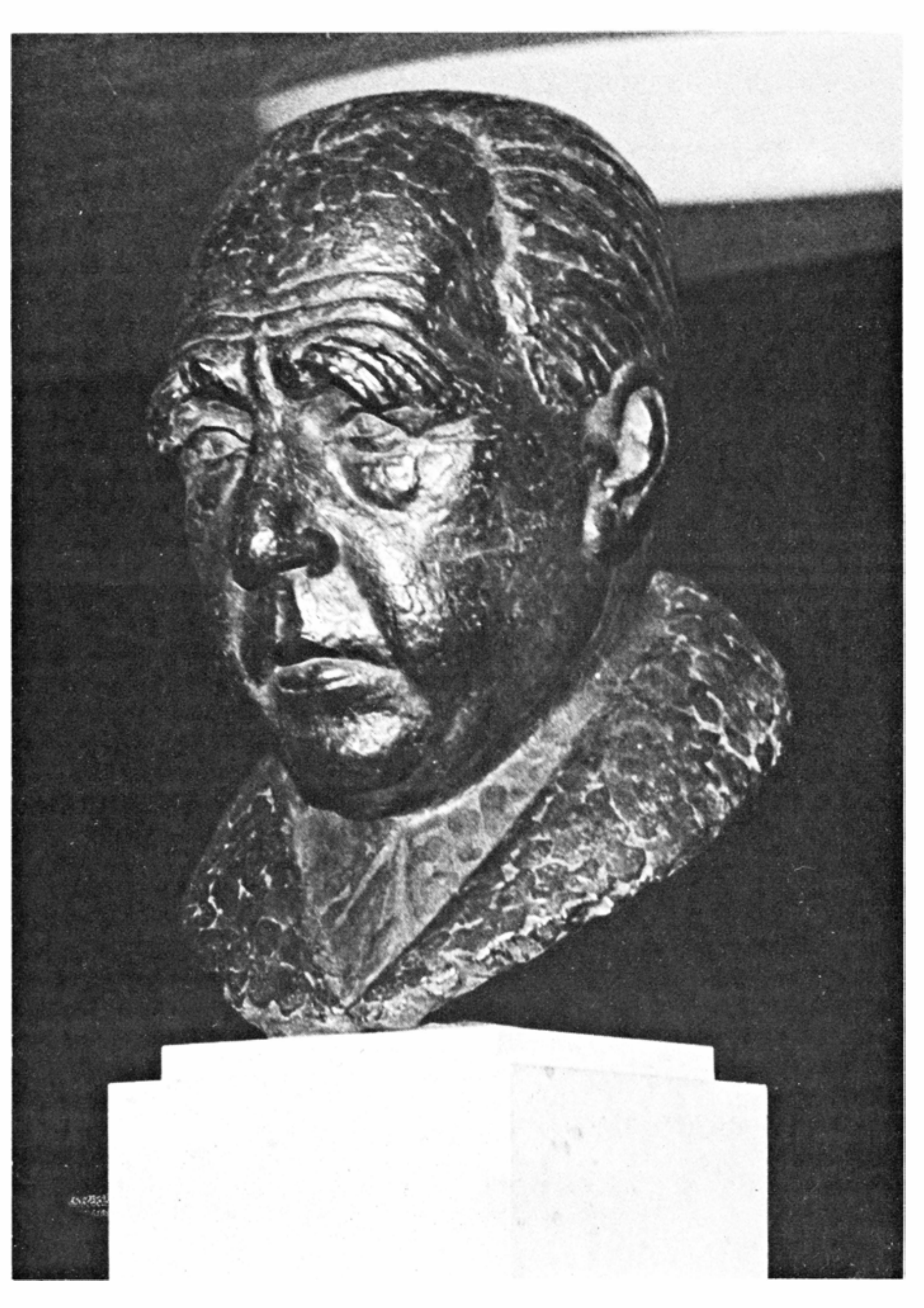
Bohr relaciona inmediatamente esta descripción complementaria del sistema cuántico con el carácter, a la vez corpuscular y ondulatorio, de los cuantos de luz y aun de las partículas materiales. Según él, "...también aquí se trata de imágenes de los fenómenos no contradictorias, sino complementarias, que solamente juntas ofrecen una generalización natural del modo de descripción clásico."

Y, comentando expresiones básicas de la mecánica ondulatoria formulada por Louis de Broglie en 1924, se complace en subrayar cómo la constante de acción de Planck aparece en ellas relacionando magnitudes de carácter corpuscular (la energía y el momento) con otras de carácter ondulatorio (el periodo y la longitud de onda, respectivamente).

En realidad, parece que esa dualidad onda-corpúsculo fue la que motivó las especulaciones de Niels Bohr, de suyo muy anteriores a las de Werner Heisenberg sobre las relaciones de incertidumbre. (Rastros de estas especulaciones aparecen en un artículo de Bohr de julio de 1925 y en una conferencia de diciembre de 1926, y están relacionadas con las experiencias de Bothe-Geiger, que imponían el carácter corpuscular del efecto fotoeléctrico.) Pero, sin duda, las discusiones entre Bohr y Heisenberg en Copenhague durante el invierno de 1926 fueron madurando las ideas de ambos. Sin embargo, es curioso que los dos pensadores necesitaran separarse para formular sus reflexiones. En efecto, fue durante las vacaciones de esquí de Bohr en Noruega cuando Heisenberg redactó su trabajo sobre las relaciones de incertidumbre, mientras que, según testimonio del propio Heisenberg, "Bohr se encontraba desarrollando los fundamentos de su idea de complementariedad".

Niels Bohr no confundirá nunca su principio de complementariedad con las relaciones de incertidumbre de Heisenberg. En su comunicación al Congreso de Como, formula así estas últimas, en lenguaje relativista: "Según la teoría cuántica, existe una relación recíproca general entre la máxima precisión con que puede definirse el vector de espacio-tiempo y el de energía-momento asociados a un mismo individuo."

Y a continuación expresa así la conexión entre esa relación de incertidumbre y su idea de complementariedad: "Esa circunstancia puede considerarse como una simple expresión simbólica de la naturaleza complementaria entre la descripción espacio-temporal y las exigencias de causalidad. Al mismo tiempo, sin embargo, el carácter general de esta relación permite hacer compatibles hasta cierto punto las leyes de conservación con la coordinación espacio-temporal de las observaciones..."



Busto conmemorativo de N. Bohr (1885-1962), que preside el vestíbulo central del Centro Europeo de Investigación Nuclear de Ginebra

Esa "expresión simbólica" alude sin duda al carácter puramente físico-matemático y no directamente epistemológico de las relaciones de Heisenberg, en oposición al principio de Bohr. Además, la complementariedad de Bohr encierra, según hemos indicado, una asimetría conceptual que no aparece en la formulación de las relaciones de incertidumbre. Niels Bohr parece querer recuperar esa asimetría conceptual, al indicar que de esos dos vectores "asociados a un mismo individuo", uno de ellos, el de energía-momento, se relaciona con las leyes de conservación (y por consiguiente con la evolución causal del sistema), mientras que el otro, el de espacio-tiempo, posibilita la observación. La compatibilidad parcial, entre ambos que formulan las relaciones de incertidumbre, "expresaría simbólicamente" no la contradicción sino la simple complemen-

tariedad de los correspondientes rasgos de la descripción física.

Wolfgang Pauli considera fundamental esta noción de complementariedad para la nueva mecánica cuántica, hasta el punto de que propone denominar a ésta "teoría de la complementariedad", en paralelismo con la "teoría de la relatividad". Pero entiende por complementariedad una relación simétrica entre dos magnitudes físicas, ligadas por las relaciones de incertidumbre, y no la relación asimétrica entre dos rasgos de la descripción física, según propone Bohr. En su famoso artículo de *Handbuch der Physik* sobre la mecánica ondulatoria, Pauli define así su noción de complementariedad, que espuriamente atribuye a Bohr: "Si la aplicación de *un cierto* concepto clásico está en una relación de exclusión respecto a *otro*, a esos dos conceptos (por ejemplo las coordenadas de posición y momento de una partícula) los llamaremos con Bohr *complementarios.*"

Esta noción que algunos llaman de "complementariedad paralela" es, en realidad, la más difundida entre los físicos actuales. Pero creemos que con ello se ha eliminado la riqueza epistemológica de la noción original de Niels Bohr.

Porque Bohr concibió su idea de complementariedad para atacar los problemas epistemológicos que plantea la física cuántica, que cree más profundos que los planteados por la física relativista y considera análogos a los problemas epistemológicos del conocimiento humano en general. Concluye así su comunicación de Como, según una reedición temprana de la misma: "Espero que la idea de complementariedad sea apropiada para caracterizar la situación [de la física atómica en relación a la física relativista], que encierra una profunda analogía con la dificultad general en la formación de las ideas humanas, inherente a la distinción entre sujeto y objeto."

No vamos a entrar aquí en las reflexiones ulteriores que en torno a la complementariedad realizó Niels Bohr en el campo general de la epistemología, de la biología o de la psicología (inspirado en ésta por W. James). Indiquemos únicamente que este principio de complementariedad, nacido de la misma experiencia atómica, ha revolucionado profundamente la antigua mentalidad mecanicista de la física, dándole un tono innegable de idealismo; que los problemas que él ataca siguen abiertos en nuestra formulación actual de la mecánica cuántica (compatibilidad del axioma de evolución con el axioma de medida). Y que cabe esperar, con Niels Bohr, que aún sea apropiado para reflexionar sobre los problemas epistemológicos de las futuras teorías cuántico-relativistas, que sin duda surgirán de nuestras actuales experiencias subatómicas. (M.G.D.)

Cadena de sucesos

Podría parecer improbable que el *status* de un pececillo tuviera graves consecuencias alimentarias, políticas y económicas para todo el mundo. Y, sin embargo, así ha ocurrido con la anchoveta peruana, *Engraulis ringins*, que hasta hace poco pululaba en vastos cardúmenes a lo largo de los 2400 kilómetros de costa del Perú, proporcionando la pesca más voluminosa de entre todas las especies marinas. Actualmente la anchoveta está

ONORANZE AD ALESSANDRO VOLTA

NEL PRIMO CENTENARIO DELLA MORTE

ATTI

DEL

CONGRESSO INTERNAZIONALE DEI FISICI

11-20 SETTEMBRE 1927 - V

COMO - PAVIA - ROMA

PUBBLICATI A CURA DEL COMITATO

VOLUME PRIMO



BOLOGNA

NICOLA ZANICHELLI

MCMXXVIII

x *Introduzione*

Segue l'elenco dei loro nomi, divisi per nazione, colle città di loro sede:

Austria: F. Ehrenhaft, Vienna; A. Smekal, Vienna.

Canadà: J. C. Mac Lennan, Toronto.

Danimarca: N. Bohr (*premio Nobel 1922*), Copenaghen; W. Heisenberg, Copenaghen.

Francia: M. Brillouin, Parigi; M. de Broglie, Parigi; P. Boucherot, Parigi; A. Cotton, Parigi; P. Janet, Parigi.

Germania: M. Born, Göttingen; J. Franck (*premio Nobel 1925*), Göttingen; W. Gerlach, Tübingen; E. Grüneisen, Marburg; M. v. Laue (*premio Nobel 1914*), Berlino; F. Paschen, Charlottenburg; W. Pauli jr., Amburgo; M. Planck (*premio Nobel 1918*), Berlino; A. Sommerfeld, Monaco; O. Stern, Amburgo; K. W. Wagner, Berlino.

India: D. M. Bose, Calcutta; M. N. Saha, Allahabad.

Inghilterra: F. W. Aston (*premio Nobel 1922*), Cambridge; W. L. Bragg (*premio Nobel 1915*), Manchester; A. S. Eddington, Cambridge; O. W. Richardson, Londra; E. Rutherford (*premio Nobel 1908*), Cambridge.

Italia: L. Amaduzzi, Parma; A. Amerio, Pavia; M. Cantone, Napoli; O. M. Corbino, Roma; E. Fermi, Roma; A. Garbasso, Firenze; G. Gianfranceschi, Roma; G. Giorgi, Cagliari; T. Levi-Civita, Roma; A. Lo Surdo, Roma; Q. Majorana, Bologna; G. Marconi (*premio Nobel 1909*), Roma; M. La Rosa, Palermo; E. Perucca, Torino; C. Somigliana, Torino; P. Straneo, Genova, V. Volterra, Roma.

Olanda: H. A. Kramers, Utrecht; H. A. Lorentz, (*premio Nobel 1902*), Leida; P. Zeeman (*premio Nobel 1902*), Amsterdam.

Russia: J. Frenkel, Leningrad; P. Lasareff, Mosca.

Spagna: E. Alcobè, Barcellona; B. Cabrera, Madrid.

Stati Uniti: A. H. Compton (*premio Nobel 1927*), Chicago; W. Duane, Boston; E. H. Hall, Cambridge; A. E. Kennelly, Cambridge; I. Langmuir, New York; R. A. Millikan (*premio Nobel 1923*), Pasadena; R. C. Tolman, Pasadena; R. W. Wood, Baltimore.

Svizzera: P. Debye, Zurigo.

Inviò la propria comunicazione, scusandosi di non poter intervenire, il Prof. P. W. Bridgman di Cambridge (U. S. A.).

Actas del Congreso de Como en honor de Alessandro Volta, y lista de participantes. Niels Bohr presentó y leyó su comunicación al mismo el 16 de septiembre de 1927

en vías de extinción, y la industria pesquera del Perú –puntal de la economía de esa nación– se ve amenazada de ruina total. La crisis parece haberse originado por los efectos conjugados de una sobreexplotación y de un cambio en el movimiento de las corrientes oceánicas.

La anchoveta –o anchoa– es conocida por muchos como adorno y adobo de pizzas y ensaladas, pero su uso comercial, más importante, y con mucho, adopta la forma de harina de pescado, complemento nutritivo para aves, vacunos y cerdos. Los industriales extraen el aceite de la anchoveta y luego la cuecen, la secan y la muelen hasta formar una harina que, en volumen, contiene un 60 por ciento de proteínas. Hasta 1972, los 10 millones de toneladas métricas de anchovetas peruanas pescadas cada año se convertían en dos millones de toneladas de harina de pescado, lo que representaba más de un tercio de la producción mundial de ese alimento y dos tercios de la cantidad del mismo que ingresaba en el comercio internacional.

Se ha estudiado intensamente cuáles fueron las condiciones insólitamente favorables que permitieron a ese pez multiplicarse en tan inmenso número. En años normales, las aguas frías de la corriente de Humboldt fluyen hacia el norte, desde la Antártida hasta las costas peruanas. Los vientos alisios del sureste empujan a la corriente mar afuera, tras lo cual es reemplazada por agua ascendente desde las profundidades, rica en nitratos y fosfatos. Estos nutrientes fertilizan el crecimiento de plancton vegetal, que las anchovetas consumen. Este pez mide tan sólo unos 10 centímetros de longitud en su edad adulta; crece rápidamente y es de vida breve. Normalmente, los descendientes del desove principal, que ocurre en septiembre, entran en la zona de pesca a principios del año siguiente, y son recogidos en cantidades significativas durante los 18 meses siguientes.

Los vientos alisios presentes en las aguas peruanas se calman periódicamente, y unas corrientes cálidas penetran al sur del ecuador. Cuando esto sucede, la población de anchovetas disminuye drásticamente por dos motivos; la carencia de aguas ascendentes fértiles retarda el crecimiento del fitoplancton, fuente nutritiva de estos peces, y el agua cálida perturba el ciclo de la reproducción de los mismos. La corriente cálida procedente del norte, y cuyo origen es todavía un misterio, es conocida como El Niño (el Niño Jesús), porque normalmente aparece alrededor de Navidad, cuando en el Hemisferio Sur es verano. El Niño suele producirse a intervalos de unos cinco años y dura unos seis meses, tras lo cual el patrón de las corrientes vuelve a su estado normal y la población de anchovetas recupera rápidamente su densidad anterior.

En 1965 ocurrió una irrupción típica de la corriente de El Niño, acompañada de un brusco descenso y subsiguiente recuperación de la población de anchovetas. Siete años más tarde, en 1972, volvieron las corrientes cálidas. Los primeros signos se observaron en el extremo norte del Perú a fines de enero; luego, la corriente de El Niño se extendió a lo largo de toda la costa peruana y permaneció allí por más tiempo que nunca (más de un año y medio). Uno de los resultados de esa situación fue que los peces se concentraron en las aguas más frías, a distancia relativamente corta de la costa, por lo que fueron especialmente vulnerables al agotamiento por pesca.

En un empeño por proteger los peces muy pequeños, nacidos el otoño precedente, el gobierno peruano declaró una veda hasta marzo de 1972. Cuando se reanudaron las faenas, se lograron buenas capturas de peces mayores (nacidos en 1970), pero pronto se hizo evidente que habían sobrevivido muy pocos peces del desove de 1971. Las capturas disminuyeron de modo impresionante en abril y mayo, conforme fueron cogidos los ejemplares mayores, y el gobierno volvió a decretar una veda en junio. A consecuencia de lo antedicho, las capturas de anchovetas, que habían alcanzado un máximo histórico de 12,4 millones de toneladas en 1970, descendieron fulminantemente a 4,4 millones de toneladas en 1972. Al año siguiente se levantó la veda por breves períodos en marzo y abril, meses en que se capturaron 1,7 millones de toneladas. Estudios realizados en ese período demostraron que la población de peces en desove seguía siendo pequeña, y que los pececillos nacidos en 1972 eran nuevamente muy escasos. En marzo de 1973, el Instituto Marino del Perú expresó su grave preocupación por la menguante población de anchovetas y recomendó una veda pesquera total hasta que se hubiese recuperado la población.

Los efectos de la súbita decadencia de la anchoveta fueron impresionantes y de largo alcance. La industria pesquera peruana en conjunto, algunos de cuyos componentes estaban demasiado endeudados, entró en bancarrota y tuvo que ser nacionalizada, formándose un monopolio estatal llamado Pesca Perú. El país perdió un cuarenta por ciento de sus ingresos en divisas y la economía peruana sufrió un grave quebranto. En Estados Unidos, Alemania Occidental y Holanda, los ganaderos y avicultores se enfrentaron a la carencia de casi la mitad de sus piensos habituales y empezaron a comprar grandes cantidades de harina y semilla de soja. La súbita demanda de soja hizo subir su precio en el mundo, lo que tuvo amplias repercusiones económicas y políticas. En Estados Unidos, la fuerte presión de compradores sobre un mercado cerealero ya escaso ayudó a poner por las nubes el precio del trigo. Desde 1973 hasta hoy, la mayor producción de soja ha hecho descender algo los precios, pero aún se sienten los efectos de la escasez de harina de pescado.

Cuando la corriente de El Niño se retiró a mediados de 1973, muchos confiaron en que la población de anchovetas recuperaría su tamaño normal, pero no ocurrió así. Las capturas anuales aumentaron ligeramente, a 3,5 millones de toneladas en 1974, y no aumentaron hasta 1976 inclusive, principalmente a causa de corrientes cálidas del Pacífico (no El Niño) que aparecieron en diversos períodos de la temporada. El año en curso, según Carlos Mendoza, agregado diplomático peruano encargado de las exportaciones de harina de pescado a Estados Unidos, se permitió pescar hasta mediados de mayo, y la captura fue de 1,2 millones de toneladas. Luego se declaró la veda a causa de la sostenida disminución de la población de anchovetas.

Nadie sabe en realidad por qué las anchovetas no recuperaron su densidad normal después de la irrupción de El Niño de 1972-73, o por qué los alevines nacidos en 1971 y 1972 fracasaron. Un informe emitido por el Comité Asesor de Expertos en Investigación de Recursos Marinos de la FAO afirma que, muy probablemente, la decadencia de las anchovetas se deba a una conjunción de factores. Unos niveles moderadamente bajos de poblaciones adultas, ocasionados por capturas intensas, pueden producir de todas maneras descendientes en número adecuado si las condiciones son normales o favorables, pero quizás una explotación pesquera desmedida, unida a una elevada temperatura del agua, haya debilitado a la población de anchovetas hasta tal punto, que no fuese capaz de recuperarse por completo. La decisión del gobierno peruano, de mantener las capturas dentro del mínimo posible durante la crisis actual, tal vez logre evitar que la anchoveta peruana se extinga como especie comercial, así como ocurrió con la sardina californiana y el arenque noruego.

Flujo de calor desde el interior de la tierra

Se puede dibujar un mapa general del flujo de calor a partir de los millares de mediciones realizadas en los continentes y en los fondos de los océanos. La distribución del flujo se interpreta dentro del marco de la tectónica de placas

Henry N. Pollack y David S. Chapman

"Sin encontrar oposición alguna entramos directamente por la lúgubre puerta abierta y empecé a mirar a mi alrededor ansioso por saber qué parte de las entrañas del infierno encerraban aquellas ardientes paredes." Así describía Dante su descenso a los infiernos. Siglos más tarde, los mineros podrían haber pensado también que aludía a su lugar de trabajo, ya que es una observación ampliamente constatada que en una mina sin una ventilación y refrigeración adecuadas hace más calor a medida que se excava en profundidad. Existen otros muchos indicios de que el interior de la tierra es caliente, siendo el más notorio el caso del volcán en erupción. Algo menos espectaculares son las zonas del globo en las que se registran manifestaciones de actividad hidrotérmica, cuyo ejemplo más palmario es el de las fuentes termales, emanaciones de vapor y géiseres del Parque Nacional de Yellowstone (Estados Unidos). Uno de los axiomas fundamentales de la física, basado en la ley de Fourier sobre la conducción del calor, dice que éste fluye desde las partes más calientes de un cuerpo hacia las más frías. Puede deducirse, por tanto, que, dado el aumento de temperatura en profundidad observado en la corteza terrestre, existe un flujo de calor hacia fuera que procede del interior de la tierra.

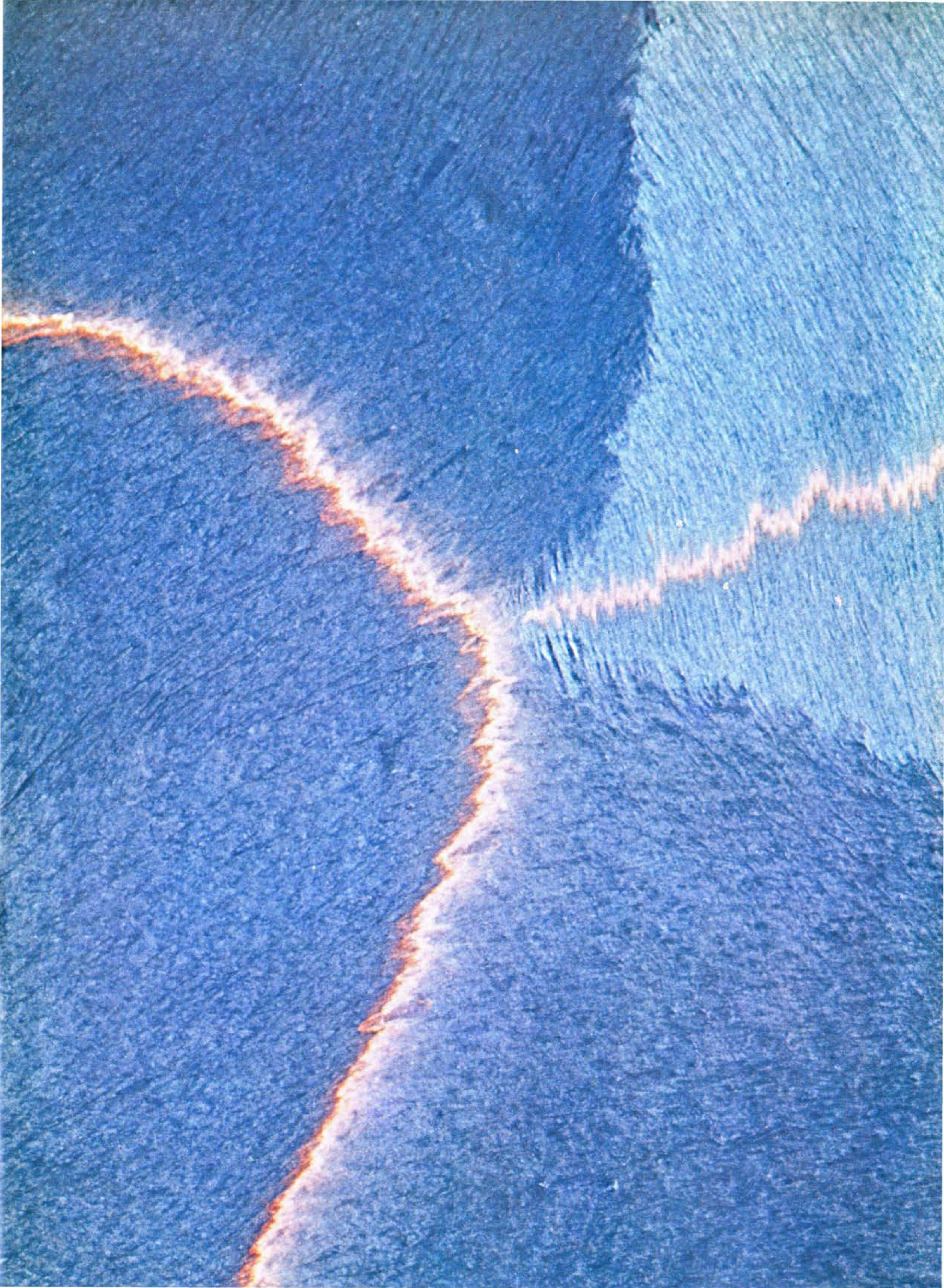
La transferencia de calor en el interior de la tierra y su paso final a la superficie por conducción a través de la corteza desempeña un papel fundamental en todas las modernas teorías de la geodinámica. En el siglo XIX, el calor interno de la tierra figuraba también de forma significativa en el prolongado debate que sobre su edad sostuvo William Thompson (lord Kelvin) con varios científicos contemporáneos suyos. El trabajo presentado por Kelvin en la Universidad de Glasgow en 1846 con el título de "Edad de la tierra y sus limitaciones determinada a partir de la distribución y movimiento del calor en su interior" fue el primero de una larga serie de artículos en los que expresaba su opinión, a saber: el gradiente térmico de la tierra (valor del aumento de temperatura con la profundidad) disminuye con el tiempo, ya que ésta se enfría desde su formación y solidificación a partir de rocas fundidas. Apoyándose en determinaciones del gradiente térmico de la tierra, obtenidas tras mediciones realizadas en minas y sondeos, mantuvo que se podía conocer el tiempo durante el cual había estado enfriándose y, por tanto, podía saberse su edad. En los diarios de las minas se podían encontrar registros de las temperaturas a diversas profundidades que Kelvin completó con sus propias mediciones hechas en Escocia. Concluyó que la temperatura, por lo menos a profundidades modestas, aumentaba entre 20 y 40 grados centígrados por cada 1000 metros de profundidad. Esta relación indicaba a Kelvin que la tierra había estado enfriándose durante sólo unas pocas decenas de millones de años, período de tiempo mucho más corto que el que muchos geólogos y biólogos de la época creían necesario para el desarrollo de los registros estratigráficos y fósiles conocidos. La controversia que esto provocó duró medio siglo y enfrentó a Kelvin contra evolucionistas tan eminentes como Charles Darwin y Thomas Huxley.

Los cálculos de Kelvin se basaban en el supuesto de que el calor perdido por la tierra procedía de la reserva térmica remanente de la etapa de fusión por la que había pasado en sus inicios. Esta suposición, que no fue controvertida en absoluto durante varios decenios, resultó ser la causa del fracaso de todos los cálculos sobre la edad de la tierra basados en las mediciones de su calor. Llegó un momento en que estas estimaciones, fundadas en el calor inicial de la tierra, no pudieron sostenerse por más tiempo; ello era debido a tres tipos de observaciones: el descubrimiento de la radiactividad por Henri Becquerel en 1896; la observación, realizada por Pierre Curie en 1903, de que la desintegración radiactiva de determinados isótopos libera calor, y la confirmación por Robert Strutt, en 1906, de que las rocas más frecuentes en la corteza terrestre contienen suficientes cantidades de isótopos radiactivos para suministrar, si no todo, por lo menos una parte importante del flujo de calor registrado en la tierra.

Medición del flujo de calor

¿Qué cantidad de calor por conducción desde su interior pierde la tierra en la actualidad? El promedio total se acerca a 0,06 vatios por metro cuadrado de superficie, es decir, unos 30 billones de vatios en todo el planeta. La cantidad de energía procedente del sol es casi 6000 veces mayor y es la responsable absoluta de la temperatura que reina en la superficie terrestre. El flujo de calor procedente del interior resulta una cantidad insignificante: el calor por conducción disipado en una superficie del

PRUEBA DRAMATICA del poder que tiene el calor interno de la tierra para moldear la geología de su superficie (fotografía tomada en 1971,en donde se observa la erupción del Mauna Ulu en el volcán Kilauea, Islas Hawai). Wendell A. Duffield, del Servicio Geológico de los Estados Unidos, sacó la fotografía desde el borde del cráter del Mauna Ulu mirando casi directamente hacia su interior. El fondo gris azulado es la corteza relativamente fría de la lava basáltica, parcialmente solidificada, que se forma en la superficie del "lago" subyacente de lava líquida a elevada temperatura. Las bandas en tierra de color naranja son grietas por las que asciende lava fundida. Este espectáculo se parece a la expansión de nuevo fondo oceánico a partir de las dorsales marinas.

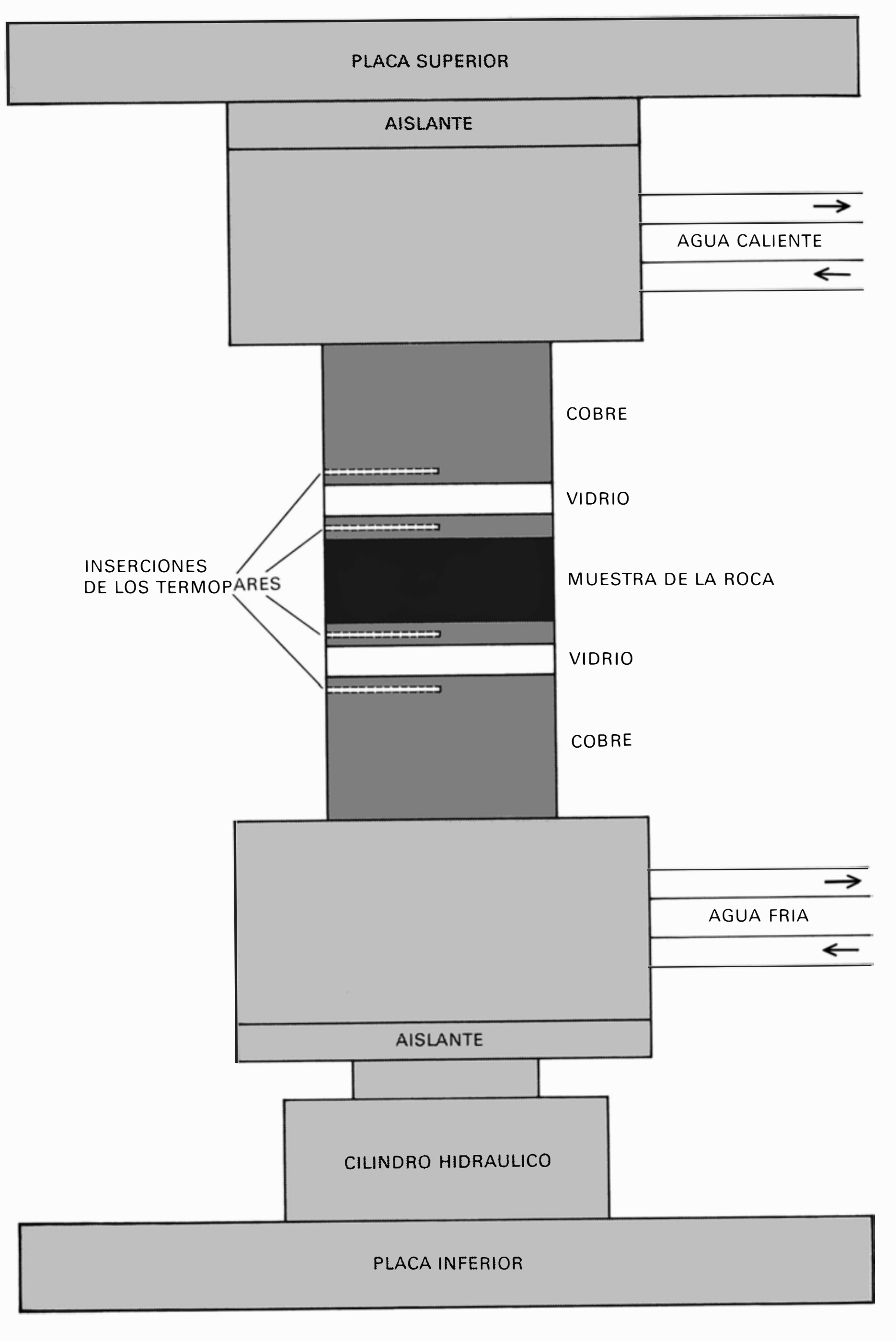




EQUIPO DE SONDEO EN ZAMBIA, dispuesto originalmente para perforar el terreno en busca de cobre, que proporcionó uno de los varios "agujeros de la oportunidad" utilizados por los autores y sus colaboradores como parte de su programa destinado a obtener medidas del flujo de calor en zonas de Africa y América del Sur en donde los datos existentes son escasos. La temperatura de la tierra se toma bajando un termómetro con resistencia eléctrica (denominado termistor) hacia el interior del sondeo y registrando mediciones a distintas profundidades. Los registros se utilizan para establecer cómo aumenta la temperatura de la roca con la profundidad; eso nos da un valor local, conocido con el nombre de gradiente geotérmico. Las mediciones de flujo de calor realizadas en Zambia se llevaron a cabo hace cuatro años, cuando los autores dejaron temporalmente la Universidad de Michigan y establecieron su base en la Universidad de Zambia. Se controlaron más de 50 de estos sondeos situados en ocho puntos geológicos distintos del Precámbrico de este país. En su opinión, los resultados obtenidos pusieron de manifiesto la presencia de materiales anormalmente calientes que se hallan situados sólo a unas pocas decenas de kilómetros por debajo de la superficie.

tamaño de un campo de fútbol equivale, aproximadamente, a la energía producida por tres lámparas de 100 vatios. Sin embargo, la evolución de la tierra ha tenido lugar a lo largo de enormes lapsos de tiempo, y una pequeña cantidad de energía a lo largo de los mismos puede ejercer un trabajo geológico importante; por ejemplo: producir la deriva de los continentes, la apertura y cierre de las cuencas oceánicas, la formación de las montañas y el desencadenamiento de movimientos sísmicos. La variación geográfica del flujo de calor procedente del interior de la tierra no es grande: la mayoría de las mediciones se encuentran dentro de un factor de tres en torno al valor medio. La distribución y características del flujo de calor en los continentes difiere de las correspondientes en las zonas oceánicas, pero la media del mismo en ambos dominios es sorprendentemente similar. Algunas áreas, como Islandia, presentan un flujo de calor extraordinario, y las zonas geotérmicas de este tipo resultan apropiadas para su utilización en la obtención de energía.

Si el calor se transmite a través de la corteza terrestre por conducción térmica, la cantidad de calor transportado es igual al producto del gradiente de temperatura por la conductividad térmica (propiedad de las rocas que define su capacidad de transmisión del calor). Cualquier estudio experimental acerca del flujo de calor de la tierra debe tener en cuenta el registro de estos dos valores. En los continentes, los gradientes de temperatura se miden bajando unos termómetros electrónicos muy sensibles (termistores) en los sondeos o midiendo la temperatura de las rocas a diferentes niveles en las minas. La perforación de un sondeo trastorna el equilibrio térmico del lugar, por eso deben dejarse transcurrir varias semanas o meses entre la perforación y las mediciones. Incluso después de que el trastorno haya perdido importancia, las temperaturas del subsuelo se ven alteradas por efectos varios: fluctuaciones diarias o anuales de la temperatura de superficie, desigualdades en la cobertera vegetal y en las formas topográficas, movimiento de las aguas subterráneas, levantamiento o erosión de la superficie exterior y variaciones del clima. La mayoría de estas anomalías disminuyen hasta niveles aceptables para profundidades superiores a unas decenas de metros; sin embargo, algunas pueden extenderse hasta varios centenares de metros. Si bien las mediciones de flujo de calor dignas de confianza pueden hacerse a veces en perforaciones poco profundas (del orden de 50 metros), los investigadores prefieren realizarlas en sondeos de una profundidad igual o superior a los 300 metros.



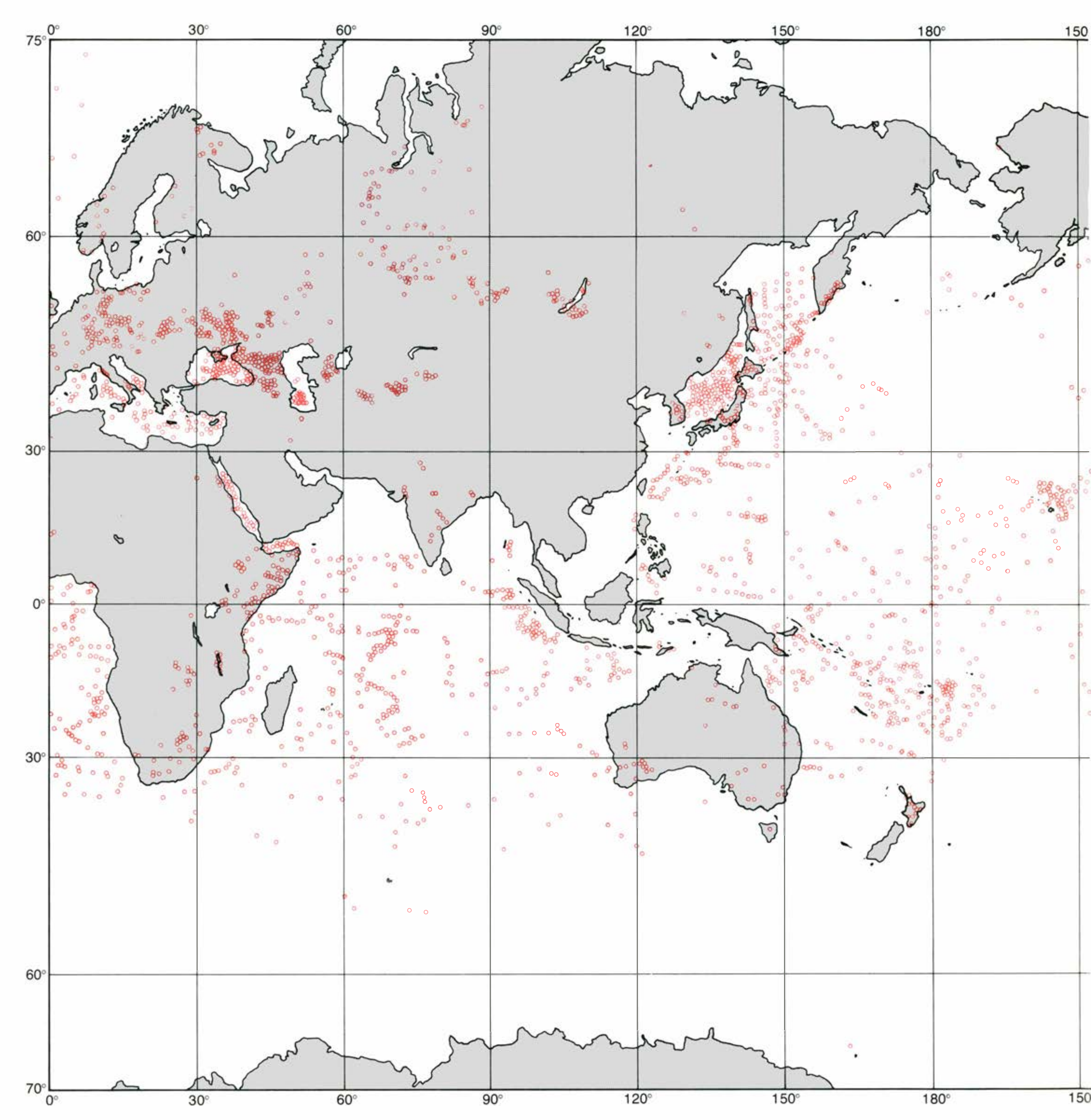
LA CONDUCTIVIDAD TERMICA de una muestra de roca dura obtenida de un sondeo se mide en un aparato cuyo esquema es el de arriba. Se corta y pule la muestra en forma de disco, colocándola en una columna entre discos de dióxido de silicio de conductividad conocida. Las diferencias de temperatura entre las terminaciones de la columna, que se mantienen constantes, producen un flujo de calor a través de la muestra y de los patrones de dióxido de silicio. La conductividad térmica de la muestra se determina midiendo el descenso relativo de temperatura a través de los componentes de la columna. El flujo de calor en el lugar del sondeo es igual al producto del gradiente de temperatura local, que se haya registrado, por la conductividad térmica de la roca.

En los fondos oceánicos, donde los sedimentos son, en comparación, blandos y la cobertura de agua marina proporciona un medio de temperatura casi constante, la perforación de un sondeo es innecesaria. Aquí los gradientes de temperatura se determinan hundiendo varios metros una sonda larga y cilíndrica en el sedimento blando y midiendo la temperatura, a intervalos de un metro, mediante termistores fijos.

Para medir la conductividad térmica suelen seguirse dos métodos. En el caso de rocas continentales duras, se corta y se pule en forma de disco una muestra del sondeo, que se intercala en una columna entre discos de dióxido de silicio

cuya conductividad es conocida. Una diferencia de temperatura constante, mantenida entre las terminaciones de la columna, produce un flujo de calor a través de la muestra y de los patrones de dióxido de silicio; el registro de la caída relativa de temperatura a través de los componentes de esa columna proporciona la conductividad térmica de la muestra. Para rocas continentales más blandas y sedimentos marinos, se inserta una aguja delgada en la muestra y se calienta en toda su longitud. A partir del registro del aumento de temperatura en relación con el tiempo, se calcula fácilmente la conductividad térmica.

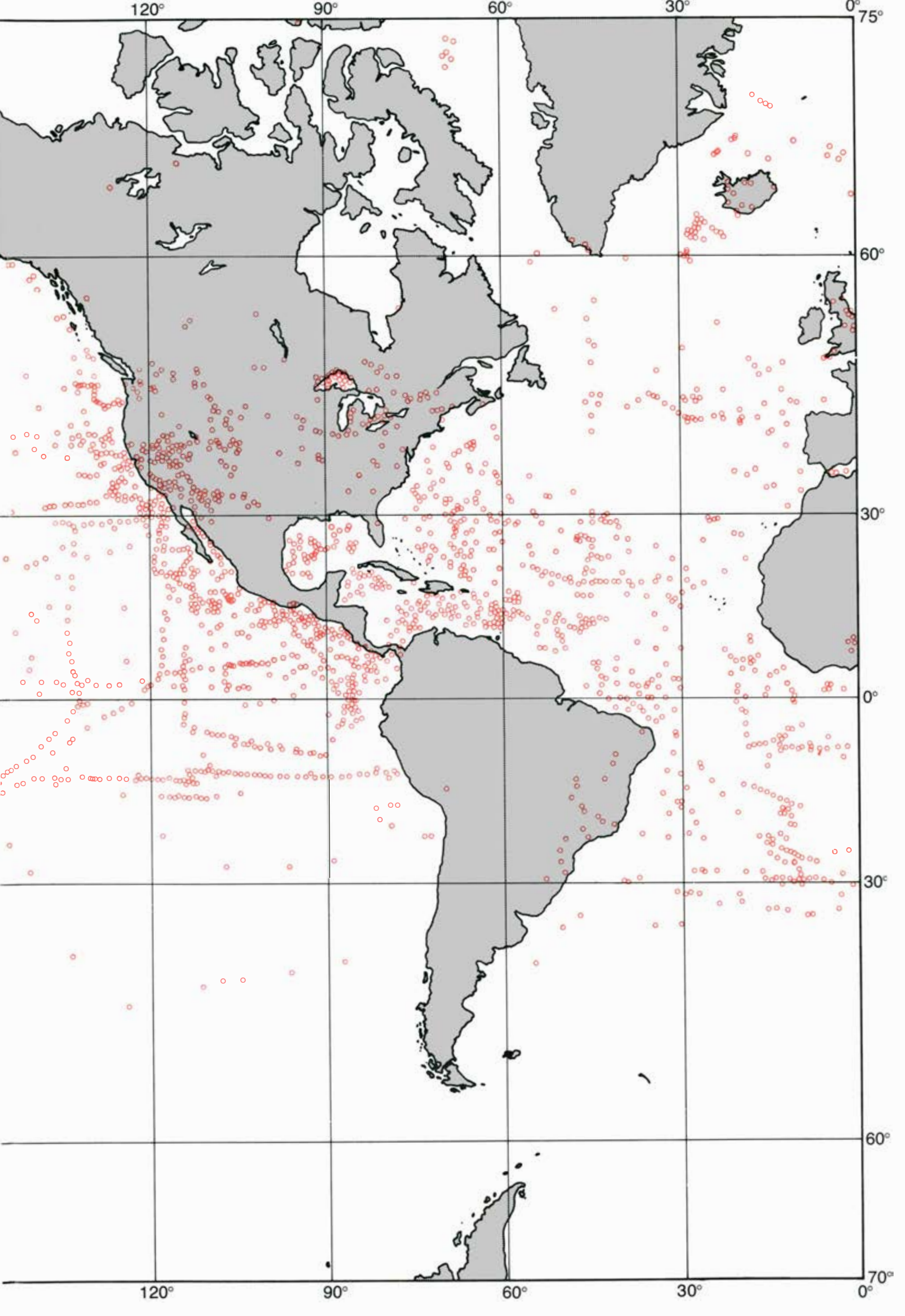
Alan M. Jessop, del Canadian Department of Energy, Mines and Resources, y John G. Sclater y Michael A. Hobart, del Instituto de Tecnología de Massachusetts, compilaron una relación en la que se incluían más de 5000 mediciones de flujo de calor de este tipo. Aunque el número de mediciones es suficiente para distintos tipos de análisis regionales, el conjunto de datos tiene una distribución geográfica desigual: los registros en el mar superan en más del doble a los obtenidos en tierra firme. Los océanos de las latitudes medias, Norteamérica, Europa y Australia, están bastante bien investigados, mientras que se carece de datos sobre extensas



DESIGUAL DISTRIBUCION de las mediciones del flujo de calor *(puntos en color)* **en zonas continentales y oceánicas. Este mapa se basa en otro preparado por el National Geophysical and Solar-Terrestrial Data Center, de los Estados Unidos. El número de estas mediciones ha**

zonas oceánicas de latitudes altas, Sudamérica, Africa, Asia y la Antártida. Durante los últimos cuatro años, nuestro equipo de la Universidad de Michigan ha contribuido, en parte, a remediar este desequilibrio geográfico en los continentes, llevando a cabo programas de mediciones de flujo de calor en Zambia, Nigeria y Brasil.

El análisis global de los datos, prescindiendo de su localización específica o de su situación geográfica, indica una dispersión muy bien definida de la distribución de los resultados, dispuestos asimétricamente en torno de un valor modal (es decir, alrededor del observado con mayor frecuencia) de 50 milivatios por metro cuadrado. Los valores singulares oscilan desde cerca de cero hasta varios centenares de milivatios por metro cuadrado, localizándose estos últimos sobre todo dentro de las zonas geotérmicas asociadas al sistema mundial de dorsales oceánicas. La subdivisión de datos en razón de su pertenencia a zonas continentales o a zonas oceánicas ofrece valores modales similares para ambos conjuntos. Sin embargo, los datos oceánicos tienen una distribución más amplia que los continentales y el alto nivel de asimetría nos da unos valores promedios (medias) para cada una de estas regiones que están muy por encima de los valores modales.



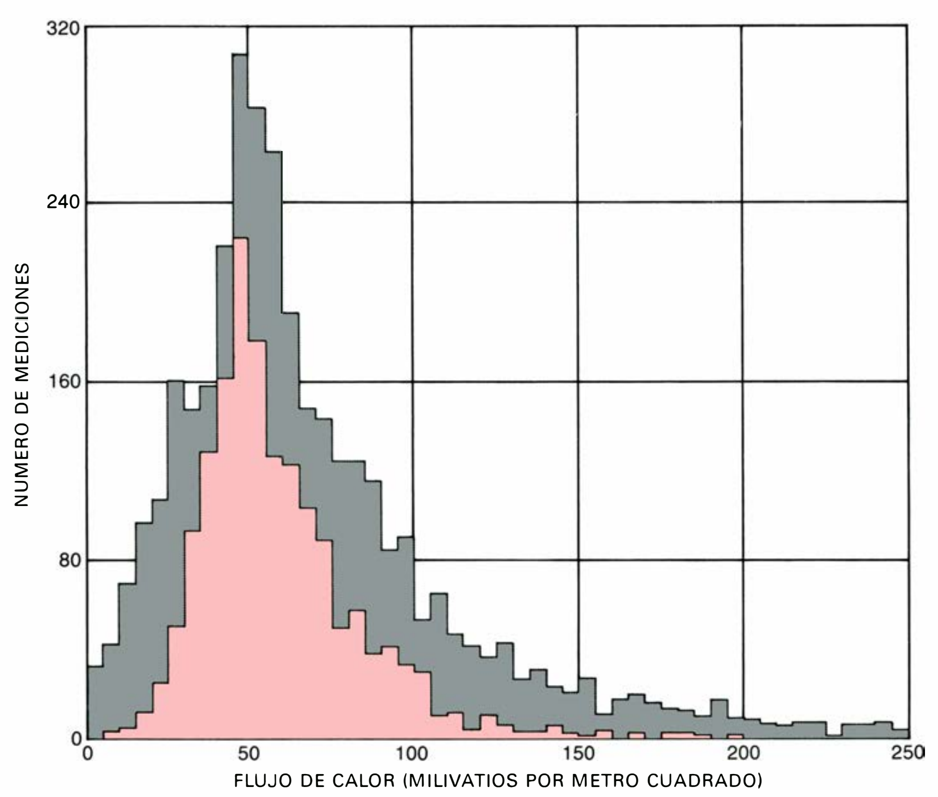
aumentando en los últimos años: de 47 en 1954 y 1162 en 1962, hasta más de 5000 en la actualidad. Las mediciones registradas en el mar representan más del doble de las realizadas en tierra firme.

Este agrupamiento global de las mediciones de flujo de calor fue de gran utilidad en el pasado, y el parecido entre las correspondientes a los continentes y a los océanos sirvió para estimular los intercambios de puntos de vista entre geofísicos y geólogos que esperaban resultados completamente diferentes de estos dos ámbitos. Sin embargo, al igual que en muchos otros aspectos de las ciencias de la tierra, las observaciones sobre el flujo de calor han encontrado forzosamente en la actualidad una nueva interpretación, al atenerse a las teorías de la expansión de los fondos oceánicos y de la tectónica de placas.

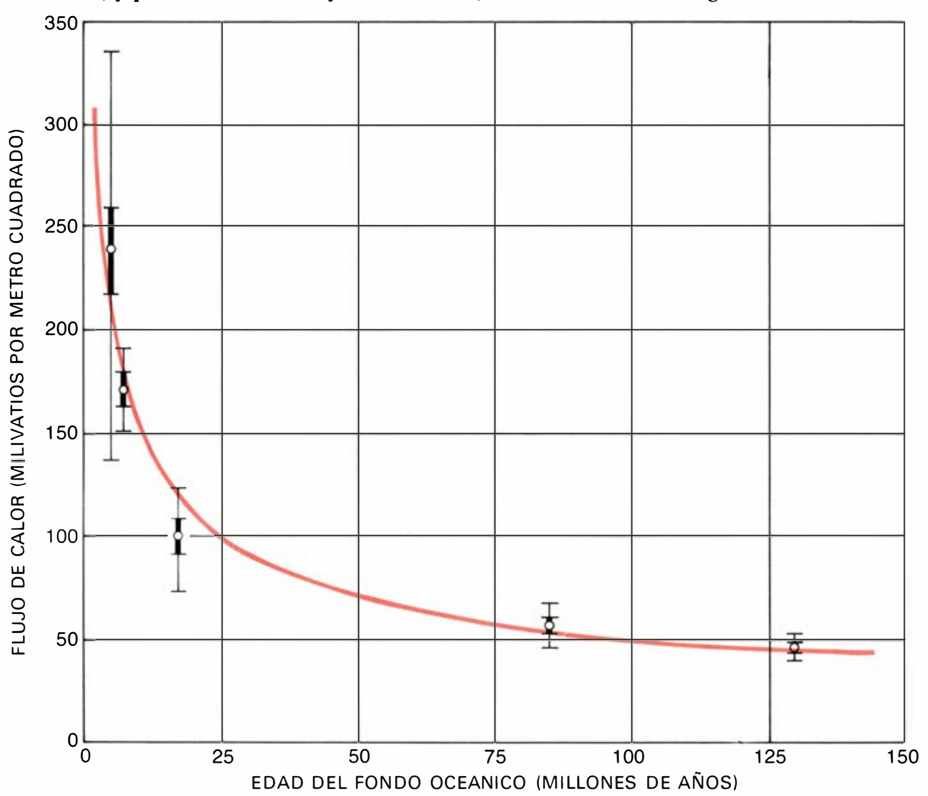
Flujo de calor y tectónica de placas

Según la teoría de la tectónica de placas, la litosfera, o envoltura más externa de la tierra, está formada aproximadamente por una docena de placas rígidas, afectadas por movimientos en la superficie terrestre. Donde las placas se están separando, el espacio que queda entre ellas se rellena de material a elevada temperatura que fluye hacia arriba procedente del manto subyacente. Este material se adiciona a los bordes de las placas que se separan, formando así las dorsales centro-oceánicas; esas rocas de nueva formación se enfrían a medida que se alejan de las propias dorsales. Al otro lado de la placa, frente a la dorsal, se encuentra generalmente una gran fosa oceánica que marca el lugar donde la placa más antigua y más fría sufre subducción, es decir, vuelve a incorporarse al interior. El calor producido por fricción y transmitido por conducción en la zona de subducción funde parte de la placa; la fracción fundida asciende a la superficie para formar los volcanes y los arcos de islas típicamente dispuestos detrás de las fosas. Estos procesos de subducción, junto con otras formas de interacción entre las placas, dan lugar al metamorfismo térmico, es decir, a la generación de magma volcánico y la formación de montañas en los continentes.

Para verificar el modelo térmico de la



UN EXTRAORDINARIO PARECIDO puede observarse en la distribución asimétrica de los valores del flujo de calor para los continentes *(color)* **y para los océanos** *(gris).* **La mayor parte de los valores registrados están comprendidos en el intervalo 20-120 milivatios por metro cuadrado, con un promedio global que está por los 60. Los valores modales, es decir, los más comúnmente observados, tanto de los continentes como de los océanos** *(picos de los perfiles),* **están más cerca de 50. Se desconoce todavía si la casi igualdad de las mediciones del flujo de calor de los continentes y de los océanos es un carácter fundamental del movimiento del calor en el interior de la tierra o es solamente una coincidencia derivada de un muestreo incompleto. Los datos de esta gráfica fueron recopilados por Alan M. Jessop, del Canadian Department of Energy Mines and Resources, y por John G. Sclater y M. A. Hobart, del Instituto de Tecnología de Massachusetts.**



EL FLUJO DE CALOR DISMINUYE conforme va aumentando la edad del fondo oceánico, tal como indica aquí la distribución de los valores del flujo medidos en cinco zonas seleccionadas del fondo del océano Pacífico. Los datos puntuales dan el valor medio de cada agrupamiento de edades. Los trazos gruesos recogen el error probable y, los finos, la desviación típica respecto de la media. Las mediciones concuerdan perfectamente con una estimación teórica del flujo de calor que cabe esperar en una placa formada por corteza oceánica que se enfría *(curva de color).*

evolución de las placas se consideran en primer lugar las oceánicas, cuya geología es relativamente simple. Edward C. Bullard, de la Universidad de Cambridge, que fue quien aportó las primeras mediciones del flujo de calor realizadas en el mar en 1954 (Atlántico), observó ya entonces la casi igualdad del flujo de calor medio en los continentes y en los fondos oceánicos. Hoy en día, con unas 3500 mediciones más, procedentes de fondos oceánicos de todas las edades, se puede apreciar una disminución sistemática del flujo de calor conforme aumenta la edad y, por tanto, la profundidad. En los sitios donde una cobertera de sedimentos espesa e impermeable impide la pérdida de calor, por circulación del agua del mar a través de la corteza oceánica fracturada, las mediciones del flujo de calor concuerdan extraordinariamente bien con las predicciones basadas en modelos matemáticos de una placa que se enfría.

Estos modelos de enfriamiento de una placa explican igualmente las características topográficas generales del fondo oceánico. Las crestas de nueva formación de las dorsales centro-oceánicas se encuentran normalmente entre 1000 y 3000 metros de profundidad, mientras que las cuencas oceánicas más antiguas están a 5500 metros. Así pues, en 200 millones de años, el fondo oceánico sufre una subsidencia de unos 3000 metros. Ello se debe a que la roca que se acaba de incorporar posee una temperatura elevada y sufre una expansión térmica, en tanto que el material más antiguo se enfría y contrae. La concordancia entre la topografía observada y la teóricamente calculada es tan notable que no parece probable explicar la subsidencia de una forma distinta de la que permite un modelo de enfriamiento simple.

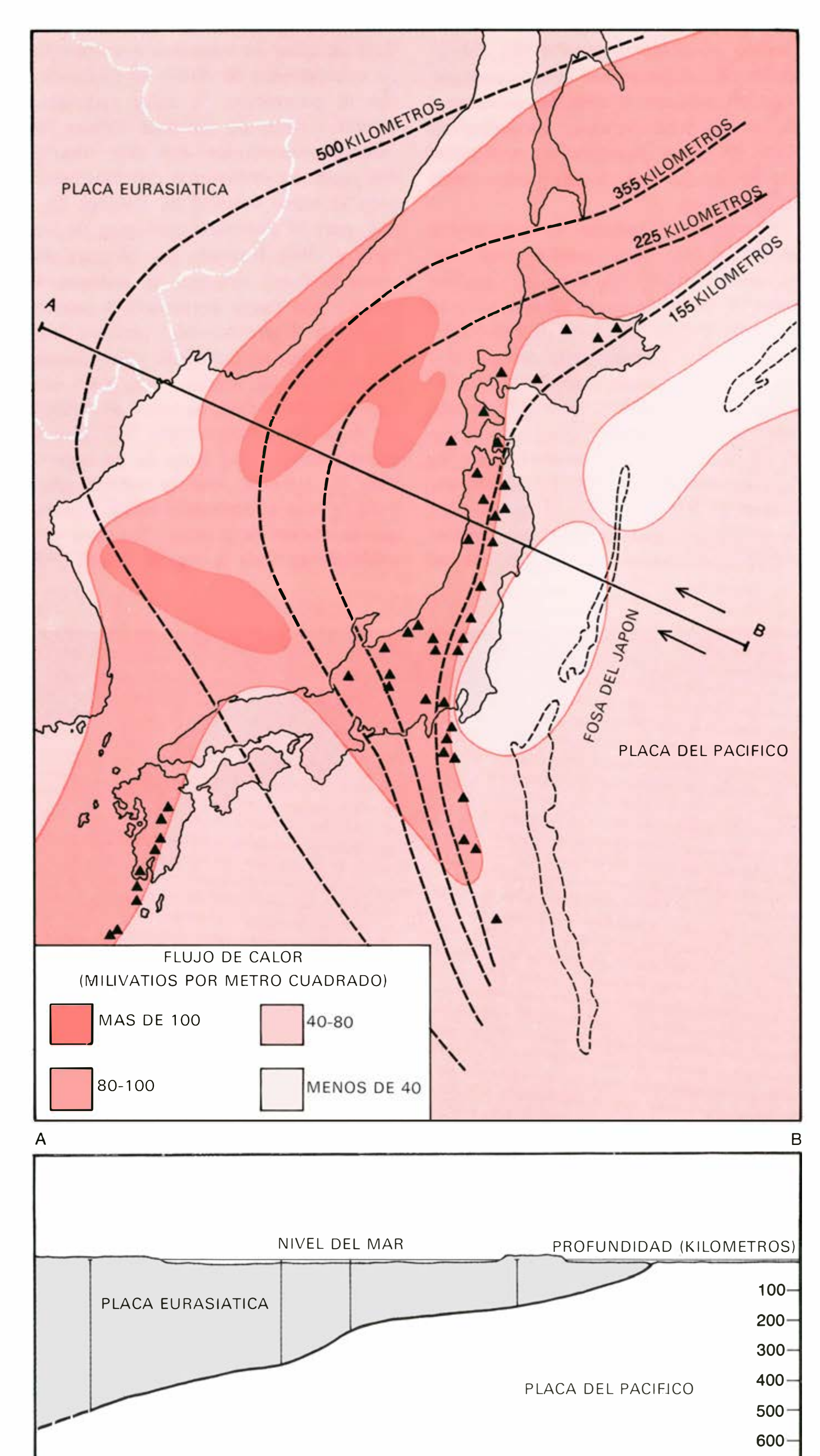
El concepto de evolución térmica de una placa oceánica puede resolvernos igualmente uno de los enigmas que desde hace largo tiempo existen en geología: ¿Cuál es la causa de las transgresiones marinas sobre los continentes? El Cretácico superior fue uno de esos períodos de tiempo en que se produjeron transgresiones marinas de importancia. Hace 100 millones de años, el nivel del mar empezó a subir, llegó a un máximo hace entre 90 y 70 millones de años, retirándose de los continentes hace unos 60 millones de años. A partir de la separación existente entre las lineaciones magnéticas del fondo oceánico puede demostrarse que el Cretácico superior fue también un período de expansión rápida del fondo oceánico. Puesto que el enfriamiento y la subsidencia de una placa

oceánica es función del tiempo, un aumento de la velocidad de expansión habría ensanchado la dorsal oceánica e incrementado su volumen. A la vez, ello habría reducido la capacidad o volumen de las cuencas oceánicas y habría desplazado el mar hacia los continentes. La regresión subsiguiente se debió obviamente a una reducción de la velocidad de expansión del fondo oceánico, que empezó hace unos 85 millones de años.

Por encima de las zonas de subducción, las características y distribución del flujo de calor son más complejas; a pesar de todo, proporcionan una información valiosa sobre el proceso de subducción. Una distribución generalmente observada en las zonas de subducción, y particularmente bien representada en el sistema de arcos de islas del Japón, es la correspondiente a un bajo flujo de calor cerca de la fosa oceánica y a un flujo muy elevado en la dirección que, respecto al arco, nos acerca al continente. La distribución sugiere que la parte superior de la placa fría en subducción actúa como elemento absorbente de calor produciendo la banda de bajo flujo observada junto a la fosa. A mayor profundidad en la zona de subducción, el calentamiento por fricción y conducción basta para fundir parte de la placa, dando como resultado el propio arco de islas volcánico y el flujo de calor más elevado existente por detrás de este último.

La región occidental de Estados Unidos ofrece otro ejemplo de zona alargada de bajo flujo de calor adyacente a una región con actividad magmática y flujo elevado. Sin embargo, en la actualidad no existe en ella ni en sus cercanías ningún área de subducción importante. Para David D. Blackwell, de la Southern Methodist University, y otros investigadores, esta zona representa una disposición y unas características de un flujo de calor fósil establecido hace unas decenas de millones de años, cuando había una subducción activa a lo largo del contacto entre la placa del Pacífico y la placa de Norteamérica.

Las características del flujo de calor pueden identificarse asimismo en ámbitos continentales que ya no forman parte de bordes de placa. Hasta ahora, se han realizado en los continentes unas 1700 mediciones, pudiéndose sacar de estos datos dos conclusiones importantes. En primer lugar, existe una disminución general en el flujo de calor que es correlativa al aumento en edad de la provincia geológica. Este resultado es afín al obtenido para los océanos, pero se ve claramente que la escala temporal difiere mucho. Mientras el flujo ca-

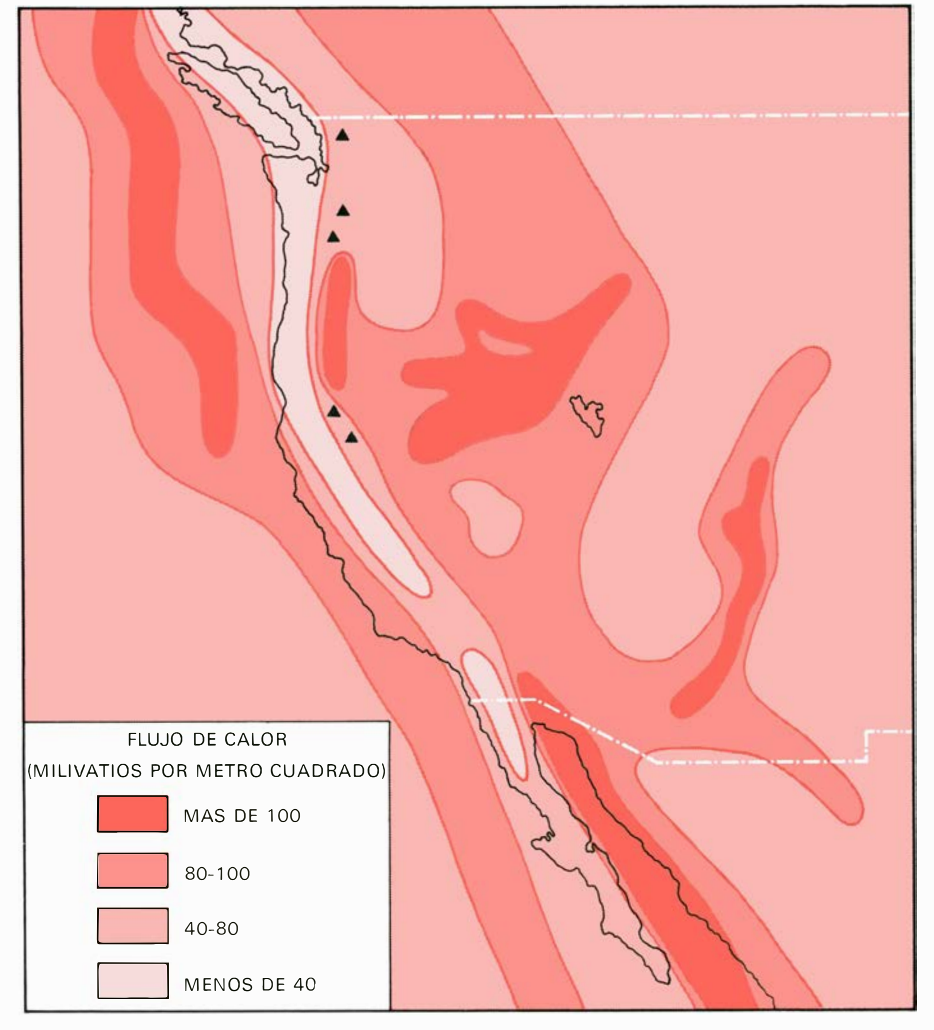


DISTRIBUCION CARACTERISTICA DEL FLUJO DE CALOR debida a la subducción de la placa tectónica situada bajo el océano Pacífico y su consiguiente hundimiento por debajo de las islas del Japón. Las flechas del mapa superior indican la convergencia relativa de las placas del Pacífico y de Eurasia; las líneas de trazos dan la profundidad de la porción de placa que sufre subducción. En el corte esquemático de la parte inferior, dibujado a escala aproximada a lo largo de la línea *AB*, se indican las profundidades correspondientes. La zona de escaso flujo de calor *(color más pálido)*, observada entre la fosa del Japón y el arco insular, sugiere que entre la superficie y una profundidad de unos 120 kilómetros la placa fría en subducción se comporta como un absorbente del calor que procede del manto terrestre. Los volcanes del arco insular *(triángulos negros)* y la región de alto flujo de calor *(color más oscuro)*, situada entre el Japón y el continente asiático, tienen su origen en el calor de fricción y la fusión parcial que se producen en la zona de subducción.

lorífico oceánico desciende por debajo de los 50 milivatios por metro cuadrado tras 100 millones de años de enfriamiento, en los continentes se encuentra un flujo de calor parecido en provincias geológicas que son cuatro o cinco veces más antiguas.

El segundo resultado de importancia estriba en que, para grandes áreas continentales, existe una estrecha relación entre el flujo de calor superficial y la radiactividad de las rocas de superficie. Que las rocas continentales, los granitos en particular, generan notables cantidades de calor debido a la desintegración espontánea de los elementos radiactivos es un hecho conocido desde comienzos de este siglo. En 1968, A. Francis Birch, Robert F. Roy y Blackwell, todos ellos de la Universidad de Harvard, demostraron que cuando las mediciones de flujo de calor se trasladan a un sistema de coordenadas en donde se relacionen con la generación de calor radiactiva debida a rocas de distintos lugares, los valores proyectados nos dan una línea recta. Se obtuvieron distintas líneas para la parte oriental de Estados Unidos, para la provincia geológica de Nevada y Utah, formada por bloques elevados y fosas rellenas de sedimentos ("basin and range province"), y para la región de Sierra Nevada, pero en cada una de ellas se mantenía una relación lineal. Lo cual implica que, para una región dada, el flujo de calor en superficie tiene dos componentes: una componente cortical, que varía de un lugar a otro de acuerdo con la radiactividad local, y otra componente más profunda, que se origina en el manto terrestre y es uniforme en toda la región. Desde 1968 se ha venido concediendo gran importancia a esta relación. Arthur H. Lachenbruch, del Servicio Geológico de Estados Unidos, ha investigado las consecuencias que ello pudiera tener en la distribución de los isótopos productores de calor existentes en la corteza terrestre y ha explicado por qué debería esperarse una disminución de la concentración de los mismos en razón exponencial con la profundidad. La variación de la componente controlada por el manto, relativa al flujo de calor observado entre provincias distintas, no ha sido tan bien estudiada. Ahora bien, basándonos en los limitados datos disponibles que poseemos, hemos deducido que, en la mayoría de las provincias caracterizadas por un determinado flujo de calor, existe una variación regular en cuanto a la procedencia del mismo, con un 40 por ciento aproximadamente del flujo superficial medio que proviene del interior de la zona de enriquecimiento cortical y un 60 por ciento que procede de abajo. Si se confirma esta procedencia selectiva, cabría pensar en que la producción media de calor de la corteza continental varía inversamente a la edad de la misma, porque, en general, las provincias más antiguas presentan un flujo de calor menor. Una relación de este tipo puede explicarse en un modelo donde la radiactividad disminuya con la profundidad, y en el supuesto de que las provincias geológicas más antiguas se hayan erosionado hasta profundidades mayores que las jóvenes.

Tom Crough, de la Universidad de Stanford, y V. M. Hamza, de la Universidad de São Paulo, aportaron otro descubrimiento interesante. Según ellos, cuando se substrae la contribución del flujo calorífico que se origina en la zona de enriquecimiento cortical, el flujo restante continúa mostrando una dependencia inversamente proporcional a la edad de la provincia, si bien la escala temporal de este flujo de calor residual es parecida a la de un proceso de enfriamiento simple, muy similar al que se observa en los océanos. Sin embargo, es obvio que el enfriamiento de las zonas continentales se encuentra en un estado mucho más avanzado y debe haber penetrado a profundidades mayores. ¿Estamos registrando el calor residual de un acontecimiento tectónico que se produjo hace 600 millones de años? Si así es, este tipo de acontecimientos deben afectar, como mínimo, a los 500 kilómetros más externos de la tierra para que cualquier tipo de calor residual llegue hoy a la superficie.

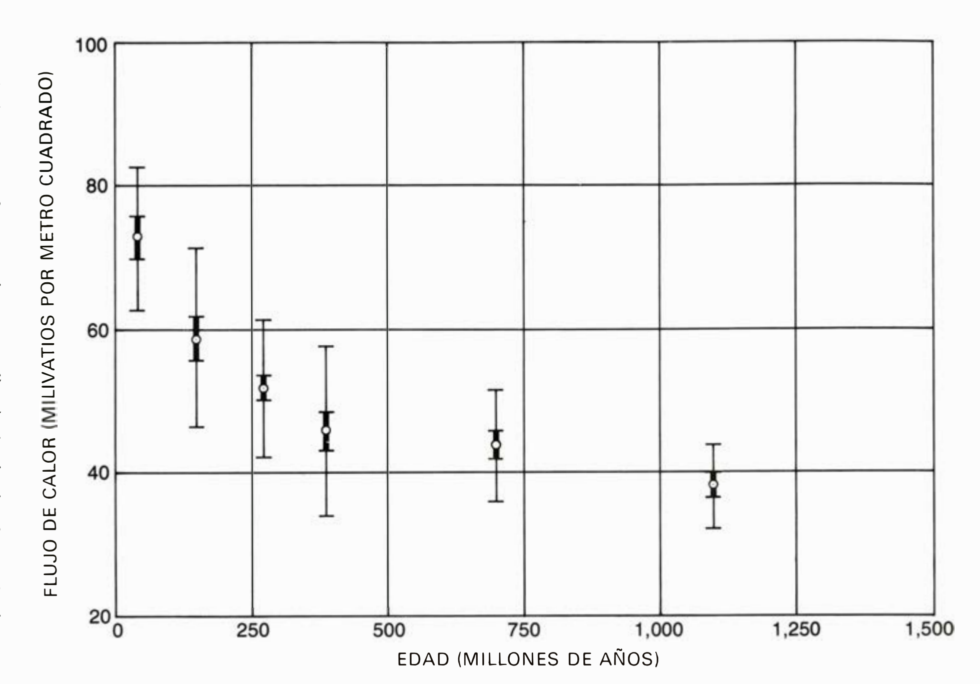


UNA ZONA DE SUBDUCCION FOSIL parece ser la responsable de la existencia de la región alargada con bajo flujo de calor que se observa junto a una región magmática de elevado flujo en la parte occidental de Norteamérica. La expansión del fondo oceánico que se produce actualmente en el golfo de California y en las dorsales de Gorda y Juan de Fuca produce un elevado flujo en las proximidades de la costa. Tierra adentro, las bandas paralelas de flujo bajo y flujo elevado marcan, respectivamente, las partes menos y más profundas de una zona de subducción que se mostró activa a principios de la era Terciaria. La subducción de las placas oceánicas situadas frente a la costa de California cesó hace más de cinco millones de años, pero persiste una distribución del flujo de calor similar a la observada en el arco japonés, en actividad. Los triángulos indican volcanes recientes.

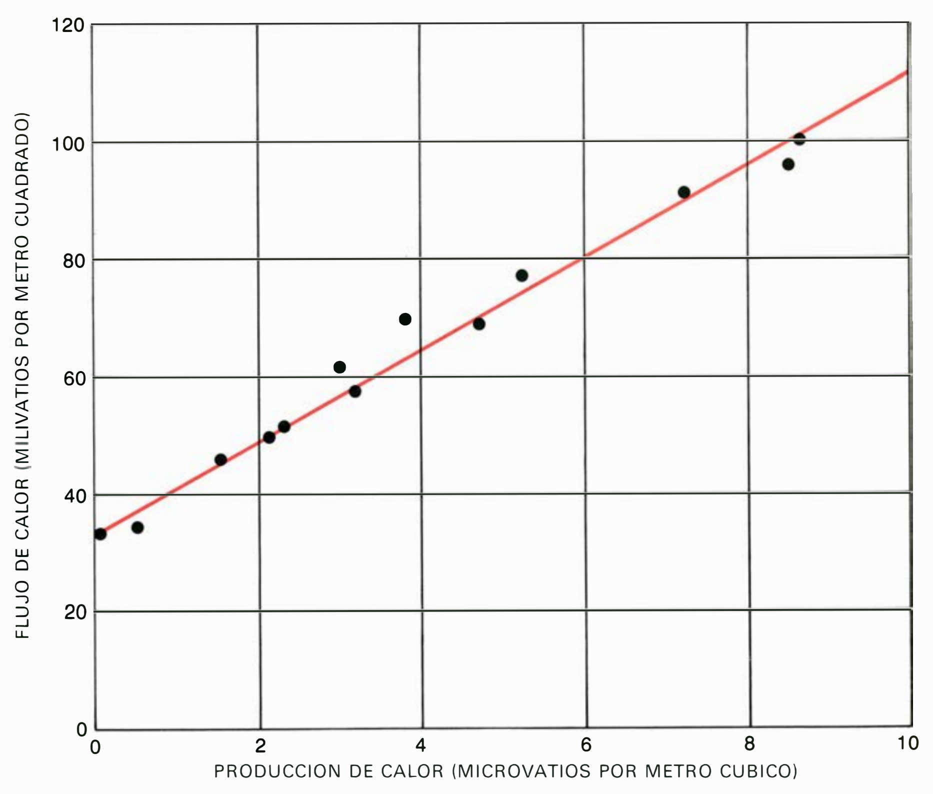
Prestemos atención, ahora, a las características generales del campo térmico de toda la tierra, con objeto de combinar los resultados de las regiones continentales y de las oceánicas. Con anterioridad a 1974 se había intentado varias veces representar gráficamente las variaciones observadas en el flujo de calor a escala global. Sin embargo, a pesar del número creciente de mediciones del flujo de calor seguían existiendo grandes zonas de la tierra de las que no se había recogido dato alguno. Se echaba de menos una representación matemática de la distribución global de las mediciones del flujo de calor para correlacionar el flujo regional con otros fenómenos geofísicos, verbigracia, el campo gravitatorio terrestre. El análisis costó muchísimo, porque había que acudir a grandes extrapolaciones para interpretar las zonas no investigadas.

Flujo de calor global

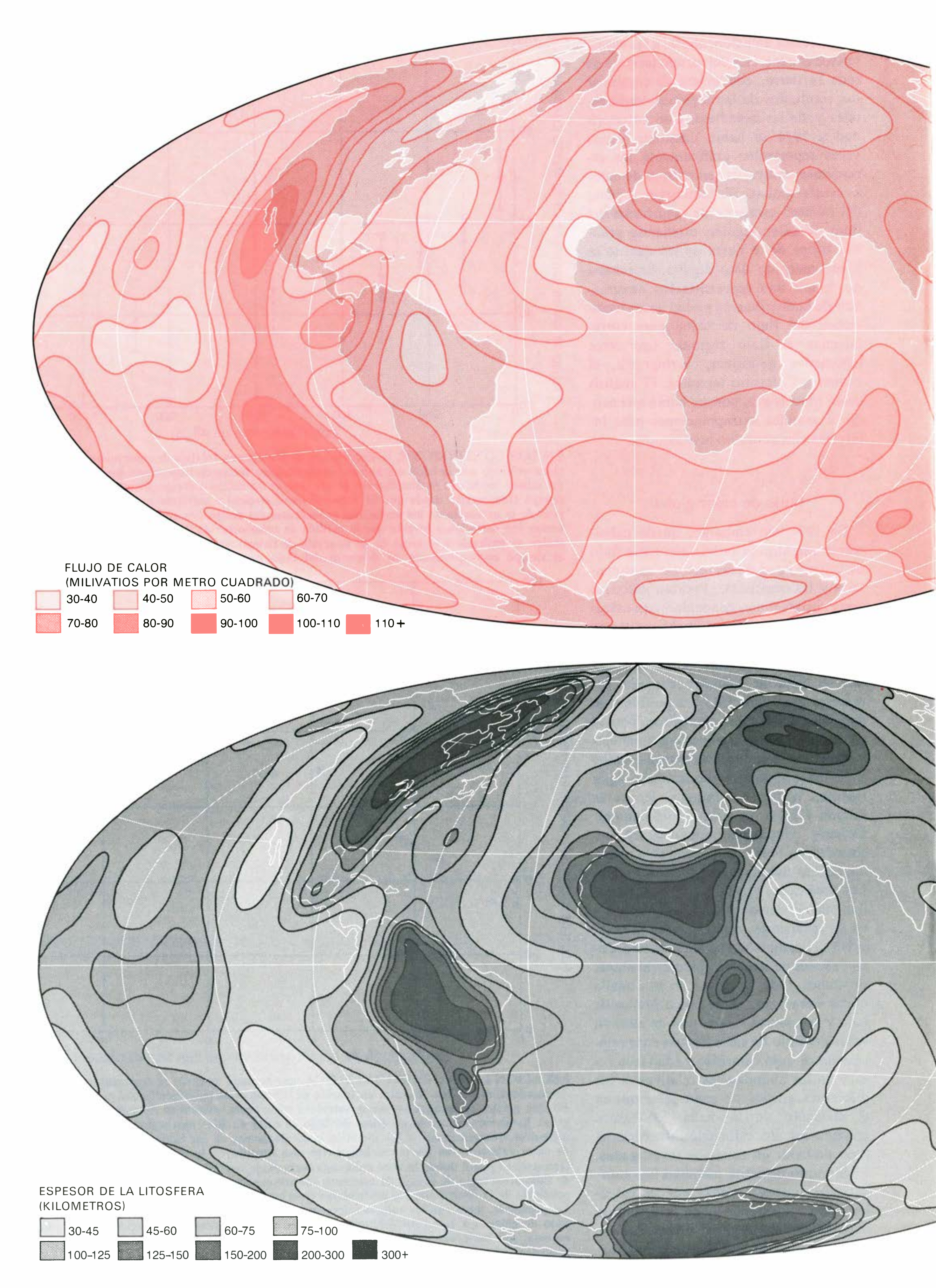
En 1974, sin embargo, quedó cabalmente establecida la relación entre flujo de calor y edad de las regiones continentales y las oceánicas. ¿Podrían utilizarse estas correlaciones conocidas para calcular el flujo de calor probable existente en zonas no investigadas, y de esta forma realizar más fácilmente, a escala mundial, el mapa de isolíneas para este parámetro geofísico? Ello era posible si se disponía de mapas geológicos con la edad de todas las regiones continentales y oceánicas. Aunque tales mapas existían para los continentes desde hacía algún tiempo, Walter C. Pitman, Roger L. Larson y Ellen M. Herron, del Lamont-Doherty Geological Observatory, no sintetizaron hasta 1974, en un solo mapa, las edades detalladas de todas las regiones oceánicas, basándose en las anomalías magnéticas de los fondos marinos y en las inversiones registradas de la polaridad del campo magnético terrestre. Poco después de haberse obtenido el mapa, dividimos toda la tierra en una malla cuyos elementos tenían cinco grados de lado y asignamos a cada uno de ellos un valor del flujo de calor basado en la relación entre flujo calorífico y edad tectónica y en las abundancias relativas de los diferentes grupos de edad presentes en el elemento de la malla. Estábamos creando así una estimación sintética del flujo de calor en zonas no investigadas. El conjunto de datos, incluidas las observaciones proporcionadas por los cálculos, podía ajustarse luego a funciones matemáticas apropiadas y representarse gráficamente con una distorsión mínima.

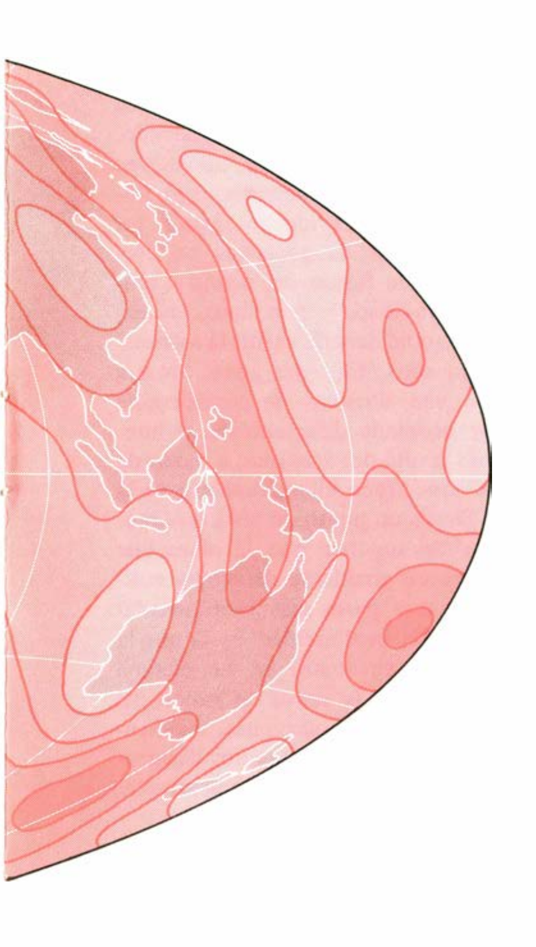


EN LOS CONTINENTES, la distribución de los valores del flujo de calor muestran también una clara relación con la edad del último acontecimiento tectónico importante que afectó a la región, del mismo modo que la distribución de los valores oceánicos correspondientes está asociada con el tiempo transcurrido desde que la roca volcánica empezó a enfriarse en una dorsal oceánica después de la extrusión de los materiales procedentes del manto. Sin embargo, en el caso de los continentes las mediciones revelan que la pérdida de calor ocurrida después de un acontecimiento tectónico se extiende a lo largo de un intervalo de tiempo que resulta cuatro o cinco veces superior al tiempo correspondiente de descenso del flujo que registra, asimismo, una placa oceánica.



RELACION LINEAL CARACTERISTICA, que se suele observar al representar, en un sistema de coordenadas, el flujo de calor en función de la producción calorífica de una región determinada. En esta gráfica se recogen los datos obtenidos en la parte oriental de Estados Unidos *(puntos negros)*. Estas variaciones en los datos del flujo de calor de diferentes regiones continentales pueden atribuirse frecuentemente a las distintas concentraciones de los isótopos radiactivos productores de calor existentes en los pocos kilómetros que constituyen la parte más externa de la corteza terrestre. El punto donde la línea coloreada corta al eje vertical del flujo de calor indica la cantidad de flujo procedente de debajo de esta zona con enriquecimiento cortical en isótopos radiactivos. La pendiente de la línea refleja la profundidad de la corteza hasta donde persiste este enriquecimiento. La pendiente, junto con los valores interceptados, sirve para caracterizar diferentes provincias de flujo de calor. La variación del flujo en el seno de una provincia está regida por la variabilidad regional del enriquecimiento en isótopos corticales. Las diferencias de flujo entre provincias proceden, sobre todo, de variaciones en la cantidad de flujo procedente de debajo de la corteza.


FLUJO DE CALOR
(MILIVATIOS POR METRO CUADRADO)
30-40
40-50
50-60
60-70
70-80
80-90
90-100
100-110
110+
ESPESOR DE LA LITOSFERA
(KILOMETROS)
30-45
45-60
60-75
75-100
100-125
125-150
150-200
200-300
300+



MAPA MUNDIAL DEL FLUJO DE CALOR realizado por los autores utilizando las observaciones disponibles complementadas con valores estimados. En él se aprecian características variables del flujo de calor a escala global. La zona principal de alto flujo de calor *(color más fuerte)* se halla en el océano Pacífico oriental, frente a las costas de América Central y Sudamérica. Esta zona coincide con la Dorsal del Pacífico oriental, donde se produce extrusión de nuevo fondo oceánico, que se aparta rápidamente del eje de la misma dando una franja relativamente ancha con gran pérdida de calor. Las dorsales de los océanos Atlántico e Indico se expanden más lentamente, produciendo, por tanto, una zona más estrecha con flujo calorífico por encima de la media. De las demás regiones con alto flujo cabe citar las cuencas oceánicas marginales del Pacífico occidental, situadas sobre zonas de subducción activas, la cordillera Occidental de Norteamérica, donde la subducción cesó hace entre cinco y diez millones de años, y la Europa alpina. Las principales regiones con flujo por debajo de la media *(color más tenue)* comprenden todas las zonas continentales formadas por escudos precámbricos y plataformas y las partes más antiguas del fondo marino.

El nuevo mapa mundial del flujo de calor, construido de esta forma, demostró, por vez primera a escala global, variaciones de este parámetro físico que habían sido detectadas en las investigaciones regionales. Todos los sistemas de las dorsales oceánicas más importantes aparecen con valores elevados del flujo de calor, ocurriendo lo propio en la Europa alpina, en la mayor parte de la región occidental de Norteamérica y en las cuencas marginales del Pacífico occidental. El centro de expansión de la isla de los Galápagos y la Dorsal de Chile destacan en el dominio de la prominente Dorsal del Pacífico oriental. Las regiones con bajo flujo de calor comprenden todos los escudos continentales y plataformas sedimentarias más importantes así como las regiones oceánicas más antiguas.

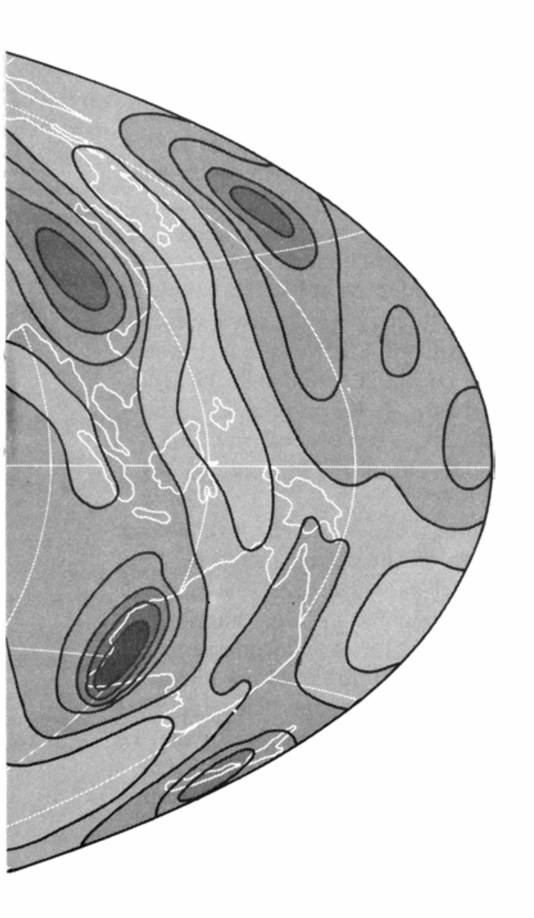
El espesor de la litosfera

Hasta hace poco, la sismología se encargaba de la determinación del espesor de la litosfera. El método sismológico consiste en observar los efectos de dispersión, de una región determinada de la tierra, en las ondas sísmicas superficiales que se propagan por ella. Los análisis de las características de esta dispersión proporcionan información sobre la elasticidad de la corteza y del manto superior pudiendo confirmar especialmente la existencia, y localizar su posición, de una zona en donde las ondas sísmicas se propagan a baja velocidad. Una zona tal pone de manifiesto probablemente la existencia de una región de fusión parcial en el manto; muchos geofísicos creen que se trata de la astenosfera, postulada en el modelo de la teoría de la tectónica de placas. Puesto que la litosfera descansa sobre la astenosfera, la profundidad de esta zona de baja velocidad equivale al espesor de la primera.

Desde hace algún tiempo, se sabe que, por debajo de los escudos continentales precámbricos antiguos, la zona de baja velocidad sísmica falta o es profunda y con escaso desarrollo. Por contra, las provincias geológicas activas y jóvenes, como las que existen en la región occidental de los Estados Unidos, tienen una zona de baja velocidad bien desarrollada y situada a poca profundidad. Recientemente han aparecido una serie de trabajos que apoyan la idea de que la profundidad de la zona de baja velocidad existente bajo los océanos aumenta igualmente de acuerdo con la edad del fondo marino, lo cual implica que la litosfera oceánica va aumentando de espesor con el tiempo.

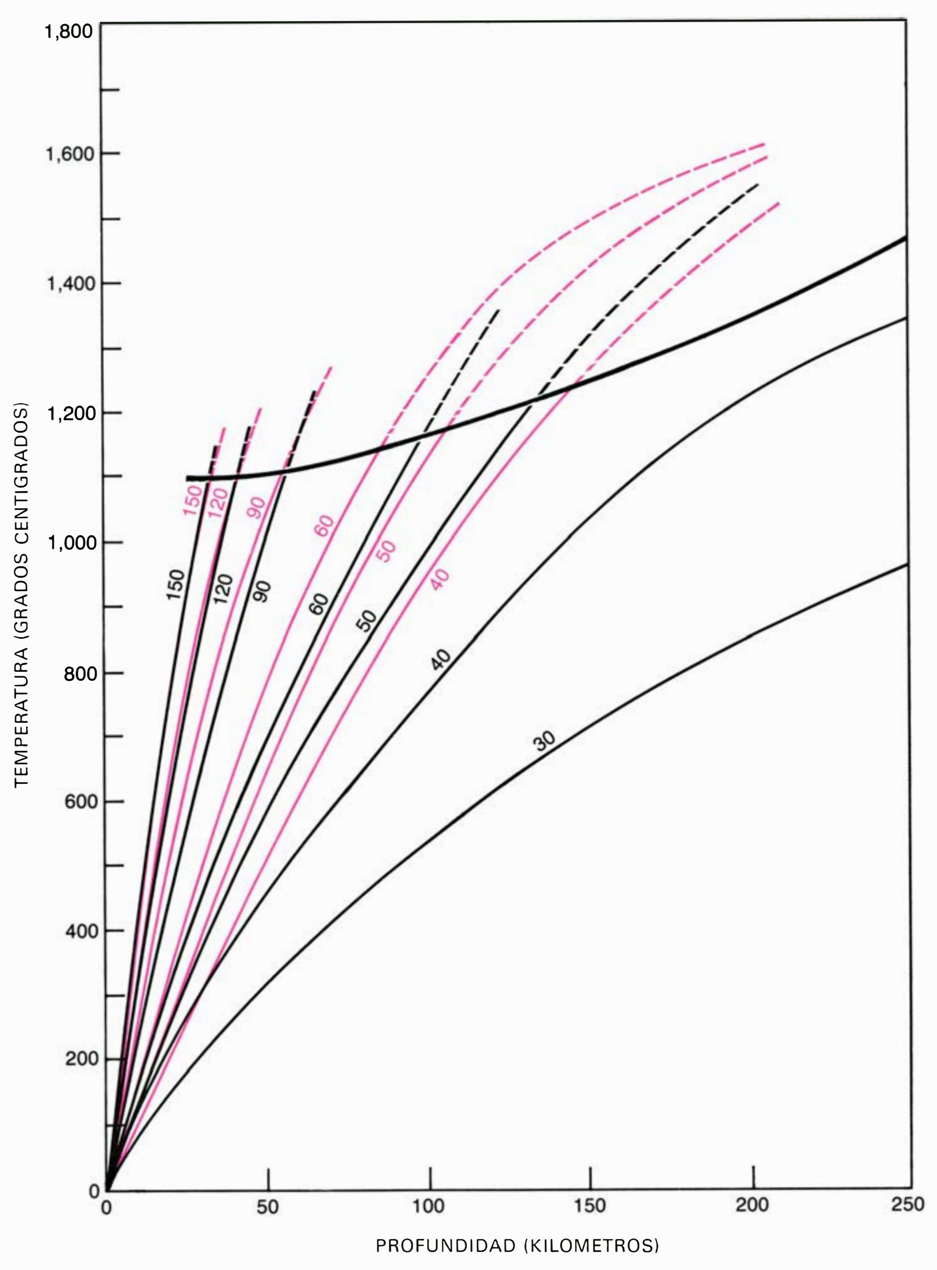
La profundidad a la cual se produce fusión parcial en el manto de una región determinada depende de la temperatura a la que la roca del manto empieza a fundir y de la variación de la temperatura en relación a la profundidad. La curva que nos relaciona la profundidad y la temperatura real, denominada geoterma, depende estrechamente, a su vez, del flujo de calor. Así pues, extrapolando mucho los datos del flujo de calor superficial pueden utilizarse para predecir el espesor de las placas tectónicas.

Puesto que, en la tierra, las mediciones directas de la temperatura se limitan a los 10 kilómetros superiores de la corteza, la extrapolación de temperaturas a profundidades del orden de los 100 kilómetros implica toda una serie de presupuestos. Se necesita saber cómo varían las propiedades térmicas de la roca en relación a la temperatura, relación entre radiactividad y profundidad y, en las regiones oceánicas, el modo de enfriamiento de la placa oceánica después



EL ESPESOR DE LA LITOSFERA TERRESTRE, o envoltura más externa y rígida de la tierra, puede estimarse seleccionando, para toda región cuadrada de cinco grados de lado, la "geoterma" que corresponde al flujo medio superficial *(véase la ilustración de la página siguiente)*. La profundidad a la que la curva geotérmica corta la incipiente curva de fusión determina el espesor de la litosfera en la región considerada. Tal como indica el mapa, la litosfera es más delgada a lo largo de las dorsales oceánicas y en otras regiones de flujo de calor elevado, y posee su mayor espesor bajo los escudos continentales. Parece indudable que este gran espesor de rocas, relativamente frías y rígidas, es el responsable de la duradera estabilidad peculiar de los escudos precámbricos.

de haberse formado en la dorsal. Recientes mediciones obtenidas en el laboratorio y observaciones de campo han proporcionado datos suficientes para la confección de modelos detallados, de suerte que se pueden calcular geotermas características con cierto grado de confianza, tanto para las regiones continentales como para las oceánicas. Las profundidades en las que se producen las condiciones de fusión parcial, predichas a partir de estos cálculos, concuerdan de modo satisfactorio con los resultados que los sismólogos han obtenido de sus estudios sobre las ondas superficiales. Las mediciones del flujo de calor y los datos sismológicos indican que las placas oceánicas aumentan de espesor a medida que crece su edad, pasando de escasos kilómetros, al poco de su formación en la dorsal, a poseer 100, o más, en las cuencas oceánicas de mayor antigüedad, en donde el flujo de calor es bajo.



DOS FAMILIAS TIPICAS de geotermas (curvas que representan la variación de la temperatura de la tierra en relación con la profundidad) para una provincia continental *(negro)* y para una provincia oceánica *(color)*. Los miembros de cada familia de geotermas poseen un número, que está de acuerdo con el valor del flujo de calor (en milivatios por metro cuadrado) producido en la superficie terrestre. Tal como cabría esperar, la temperatura que se encuentra a una determinada profundidad en una región con elevado flujo de calor será más alta que la correspondiente a la misma profundidad en otra región de bajo flujo. La curva negra de trazo grueso, cortada por la mayoría de curvas geotérmicas, representa la temperatura a la que las rocas empezarán a fundirse en el manto terrestre. La profundidad a la que se observa esta fusión es variable, y depende del flujo de calor y de la geoterma de la región considerada. Por debajo de algunas regiones continentales con escaso flujo de calor no se produce fusión probablemente. Muchos geólogos y geofísicos creen que la base de la litosfera rígida viene definida por el inicio de la fusión parcial.

Las partes continentales de las placas tectónicas muestran igualmente una variación sistemática de espesor, desde 40 kilómetros, en provincias geológicas jóvenes cuyo flujo de calor es elevado, hasta varios centenares de kilómetros bajo los escudos continentales, cuyo flujo característico es mucho menor. En algunos escudos, la geoterma no corta la curva de fusión del manto a ninguna profundidad, y, por tanto, en sentido estricto no debería existir la astenosfera bajo ellos. En estas zonas, podría hallarse una litosfera de considerable espesor acoplada directamente al interior más profundo, actuando a modo de "ancla" que frena el movimiento de la placa. Desde un punto de vista más realista, es de suponer que la placa que constituye un escudo se desacopla realmente de esta parte más profunda, tal vez a aquella profundidad en la que la geoterma se encuentra en su máxima proximidad con la curva de fusión del manto. De algún modo, pues, los escudos continentales se hallan hoy desarrollando una actividad parecida al arrastre de un ancla.

Historia térmica de la tierra

¿Qué puede decirse acerca del estado térmico de la tierra en tiempos remotos del pasado? Al tratar de este tema, la mayoría de los autores empiezan explicando por qué este problema es difícil, para refutar luego una serie de posibles soluciones en el caso de que alguno de los supuestos previos resulte no ser válido. (¡La lección de Kelvin no ha caído en saco roto para otros!) Sin embargo, cabe avanzar algunas observaciones de las que se pueden deducir ciertas conclusiones generales, en el bien entendido de que los detalles permanecen necesariamente confusos. En primer lugar, se debe tener en cuenta la abundancia, en la corteza terrestre, de los principales isótopos radiactivos que son generadores de calor: torio 232, uranio 238, potasio 40 y uranio 235.

La corteza continental tiene un espesor medio de 40 kilómetros (menos del 1 por ciento del radio terrestre); a pesar de ello, su contenido en esos isótopos productores de calor es suficientemente elevado como para que el 40 por ciento del flujo de calor de la superficie terrestre se genere dentro de la propia corteza. La concentración de isótopos en la corteza oceánica es menor pero, con todo, representa un enriquecimiento importante. La consecuencia de esta concentración hacia las partes superiores es que ha habido una segregación geoquímica de

primer orden en el interior de la tierra. Que las rocas continentales de una edad superior a los 3500 millones de años muestren tal enriquecimiento indica que la segregación tuvo lugar en una de las primeras etapas de la evolución terrestre, y muy probablemente al mismo tiempo que la tierra se diferenciaba dando un núcleo metálico denso y un manto de silicatos más ligero.

La consideración de cómo determinadas propiedades físicas de la tierra, por ejemplo, la resistencia a la ruptura y la viscosidad, cambian con la temperatura nos aporta una valiosa información sobre los procesos térmicos y tectónicos del interior de la misma. La viscosidad mide la capacidad que tienen los materiales (sólidos inclusive) para fluir. Un material muy viscoso es casi rígido, mientras que otro que lo sea poco resulta muy afín a un fluido. Las temperaturas elevadas determinan generalmente una viscosidad más baja. En la superficie terrestre y en el interior de la litosfera, las rocas son relativamente frías y rígidas pero, a más profundidad, el aumento de temperatura correlativo a esta última va acompañado casi con toda seguridad de un descenso de viscosidad, lo cual permite, en definitiva, que el interior se comporte como un fluido. En consecuencia, resulta probable que en ese interior dominen los movimientos de tipo fluido causados por diferencias de densidad, cuyo origen hay que buscarlo en sus propiedades térmicas y en sus diferencias de composición, contrastando esto último con el régimen térmico, únicamente producido por conducción, que existe en el interior de la litosfera.

En efecto, durante los primeros estadios de la evolución de la tierra, su disminución gradual de viscosidad, a medida que se iba calentando en presencia del calor radiactivo, fue probablemente responsable del inicio del proceso de estratificación por densidades, que dio lugar al núcleo y a una concentración hacia arriba de los isótopos productores de calor. La reestructuración de la masa de la tierra, una vez formado el núcleo, liberó energía gravitatoria, con lo cual debió acelerarse el proceso. La formación del núcleo de la tierra fue un acontecimiento mecánico y térmico, que se produjo una sola vez a lo largo de su historia, y que no puede parangonarse, en cuanto a alcance o intensidad, con los acontecimientos que ocurrieron en períodos posteriores, a pesar del interés actual, en el campo de las ciencias de la tierra, por la geodinámica contemporánea encuadrada en el marco de la teoría de la tectónica de placas.

La larga evolución térmica de la tierra está estrechamente ligada a la abundancia y duración de sus isótopos productores de calor. Para que un isótopo ocupe un puesto significativo en la historia térmica de la tierra debe ser suficientemente abundante y poseer una vida media lo bastante larga como para proporcionar cantidades importantes de calor durante un tiempo comparable a la edad del planeta (4600 millones de años). Sólo los isótopos arriba mencionados (torio 232, uranio 238, potasio 40 y uranio 235), con sus respectivas vidas medias de 14.100, 4510, 1260 y 710 millones de años, cumplen evidentemente estas condiciones.

Si se tienen en cuenta sus abundancias relativas y se calcula la cantidad de calor que se generaba hace 3000 millones de años, encontramos que entonces se producía 2,2 veces más calor por desintegración radiactiva que en la actualidad. Esta mayor producción de calor se reflejaba probablemente en un aumento proporcionado del flujo de calor superficial; de ello puede deducirse que, entonces, las placas litosféricas eran probablemente más delgadas, se fracturaban con mayor facilidad y, por tanto, su extensión era menor, y mayor su número en comparación con las placas actuales. La astenosfera subyacente a las placas se encontraba en un estado de actividad más vigoroso que el que manifiesta hoy en día.

En el futuro, la litosfera continuará engrosándose mientras que la astenosfera se hará más viscosa debido al continuo enfriamiento de la tierra y a la lenta desintegración de sus fuentes radiactivas de calor. El movimiento de esas placas, con un espesor cada vez mayor, se volverá más lento y retardado, si bien podrían anticiparse interrupciones en esa larga evolución, parecidas a la fragmentación y dispersión del "supercontinente" Pangea a lo largo de los últimos 180 millones de años.

Conforme los escudos continentales continúen engrosándose y comportándose a modo de importantes anclas viscosas, se podrá esperar que el movimiento de las placas cese finalmente desembocando en la terminación de la fase de tectónica de placas por parte de la evolución terrestre. Así pues, el grupo cada vez más restringido de científicos que se dedican al estudio de la tierra, y que aún mantienen puntos de vista inmovilistas, pueden tener el pequeño consuelo de que, al final, la tierra se ajustará a su modelo. Sin embargo, deben armarse de paciencia, ya que esto ocurrirá probablemente de aquí a dos mil millones de años.

Canguros

Estos marsupiales saltadores han evolucionado relativamente aislados durante unos 25 millones de años. Sus estrategias de adaptación son muy afines a las de los ungulados de las sabanas semiáridas del Viejo Mundo

T. J. Dawson

Los canguros poseen cierto poder de fascinación porque se apartan un tanto de la noción habitual que se tiene de un mamífero. Difieren de otros grandes mamíferos porque crían a sus pequeños en una bolsa y porque saltan. La explicación que a menudo se da de la singularidad de estas características es la de que los canguros son marsupiales, y, éstos, mamíferos primitivos. No se trata de ninguna respuesta, sino que refleja la impresión de las gentes que viven en el hemisferio boreal, impresión que parece desprenderse de ciertas afirmaciones a propósito de la zarigüeya de Virginia, único marsupial de América del Norte. Con frecuencia se describe a la zarigüeya como mamífero arcaico y primitivo, que apenas si ha cambiado desde la era de los dinosaurios. Tal afirmación acaso posea cierta base, ya que se han encontrado fósiles de marsupiales de 100 millones de años de edad, contemporáneos de los dinosaurios de finales del Mesozoico, los cuales se han clasificado dentro del mismo género *(Didelphis)* que la zarigüeya de Virginia. Ahora bien, aun admitiendo que en muchos sentidos la zarigüeya pueda clasificarse como un fósil viviente, no puede decirse lo mismo de los canguros. Estos marsupiales australianos representan una especialización nueva y afortunada dentro de un mundo en proceso de cambio.

La familia Macropodidae (literalmente "grandes pies") reúne unas 45 especies, entre las que se cuentan no sólo las cinco o seis especies de grandes canguros, sino también otros parientes menores, igualmente saltadores, que abarcan desde los wallabies, que siguen en corpulencia a las especies antes citadas, hasta las ratas canguro, del tamaño de un conejo. El grupo evolucionó a partir de animales ramoneadores de bosque hasta las especies actuales, que pastan predominantemente en zonas abiertas. Este proceso constituye una radiación adaptativa, ubicada en Australia, que puede compararse a las radiaciones de mamíferos no marsupiales ocurridas en el resto del mundo. El éxito de los macropódidos, posiblemente sin comparación con el de otros marsupiales australianos, estuvo quizá relacionado con el ensanchamiento, durante el Mioceno, del área ocupada por la sabana en el interior de Australia, hace entre 10 y 25 millones de años.

Durante esa era se registraron acontecimientos similares en otras partes del mundo. Fue una época de expansión para Artiodactyla, orden de mamíferos ungulados con un número par de dedos. Sabemos bastante sobre la biología de los artiodáctilos, en especial sobre la de los pertenecientes al suborden de los rumiantes, pues en él figuran algunos animales domésticos, como la oveja, la cabra y la vaca. Las adaptaciones visibles en los canguros representan soluciones a los problemas que plantea la supervivencia en un medio ambiente como el de la sabana, soluciones que difieren de las asumidas por los artiodáctilos. Además, aun en los casos de mayor similitud, dichas soluciones presentan diferencias de detalle con respecto a los no marsupiales. Un punto importante de esta cuestión es cómo se compagina el supuesto primitivismo de los marsupiales con las referidas singularidades evolutivas y de adaptación a la vida en la sabana.

Un breve repaso de la evolución de los mamíferos nos ayudará a establecer las relaciones entre los marsupiales y otros miembros de dicho grupo. Se cree que todos los mamíferos descienden de un reptil mamiferoide que vivió a finales del Triásico, hace más de 200 millones de años. Poco después, alrededor de 180 millones de años atrás, los troncos que llevarían a los dos principales grupos de mamíferos se separaron. Uno de los grupos es el de los Prototerios, compuesto por los mamíferos ponedores de huevos, del orden Monotremata; entre los representantes actuales de este orden se encuentran el ornitorrinco y el equidna australianos. El otro grupo es el de los terios, formado por mamíferos marsupiales y mamíferos placentados, estos últimos caracterizados por la posesión de una verdadera placenta, con corion y alantoides. Así, pues, los marsupiales están más estrechamente emparentados con los placentados que con los monotremas, los cuales conservan muchos rasgos reptilianos aparte de la puesta.

La radiación inicial de los terios a comienzos del Cretáceo se desplegó a partir de minúsculos insectívoros y se apoyaba, indudablemente, en adaptaciones a las plantas con flores que acababan de desarrollarse y a los insectos polinizadores que acudían a ellas. El punto de bifurcación entre marsupiales y placentados se sitúa durante este período, hace unos 130 millones de años. Los datos proporcionados por diversas fuentes de información sugieren que los marsupiales se desarrollaron en América del Norte, y fueron el grupo de terios que prevaleció en esta región del globo durante la mayor parte del Cretáceo, hasta hace unos 70 millones de años.

Por lo que parece, los placentados evolucionaron inicialmente en Asia, y sólo llegaron a América del Norte a finales del Cretáceo, tras la desaparición de la barrera oceánica que separaba los dos continentes. A esta invasión de América del Norte siguió una radiación evolutiva de los placentados, fenómeno que coincidió con la extinción de todas las líneas de marsupiales, anteriormente tan abundantes, a excepción de una. Se ha apuntado que la causa del aniquilamiento de los grandes marsupiales debió ser la competencia instaurada por los placentados invasores, pero hay otras causas que pueden ser igualmente responsables, pues ésta fue también la época del eclipse definitivo de los dinosaurios.

El seguir los pasos de la evolución, desde los diminutos marsupiales insectívoros del Cretáceo de América del Norte hasta los canguros de la sabana y de los bosques abiertos de Australia, ha supuesto desde hace tiempo un problema para el científico. No obstante, la aparición de pruebas evidentes en favor de la teoría de la deriva continental y de la tectónica de placas ha suministrado numerosos indicios acerca de cómo penetraron los marsupiales en Australia. Parece ser que, aun cuando a fines del Cretáceo los marsupiales estaban desapareciendo de América del Norte, por esa época ya se habían extendido ampliamente por América del Sur. Desde allí se desparramaron en una nueva radiación. Verosímilmente, los marsupiales australianos se originaron a partir de esta fauna, en un período en el que Sudamérica y Australia se encontraban unidas a través del continente antártico. El tiempo transcurrido desde esta ulterior expansión puede cifrarse, como máximo, en unos 50 millones de años, pues a partir de entonces Australia y la Antártida se separaron.

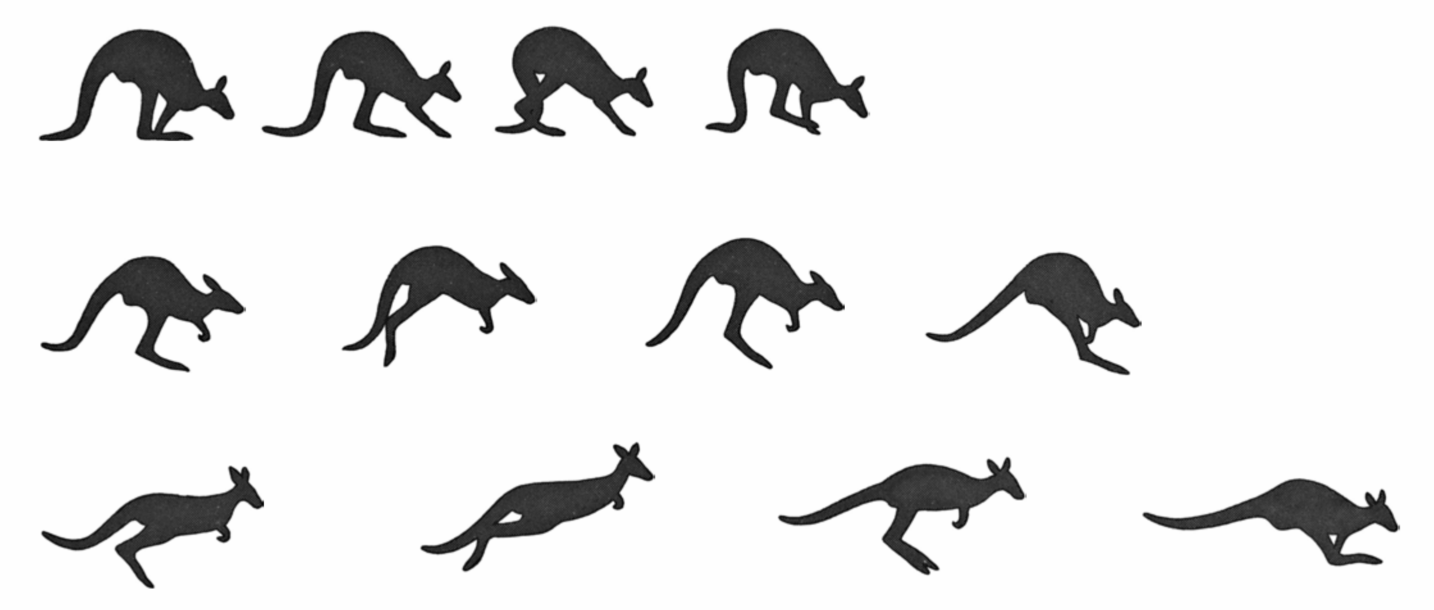
Son pobres las colecciones de fósiles de este período recogidas en Australia, por lo que no poseemos información directa sobre los primeros marsupiales que emigraron a este continente. Sin embargo, Michael Archer, del Queensland Museum, ha expresado la opinión de que arribaron en sucesivas oleadas. Su propuesta se basa en la similitud existente entre los dientes de los marsupiales australianos y los de los sudamericanos, tanto vivientes como extinguidos. Cualesquiera que fueran las formas adoptadas hace 25 millones de años por los primeros marsupiales australianos del Mioceno superior, el proceso de diferenciación había establecido ya sus grupos principales. Uno de estos grupos era el de los macropódidos, de los cuales el fósil más primitivo que se conoce es una pequeña rata canguro encontrada en terrenos miocénicos de la región del lago Eyre.

Entre los aspectos más significativos del cambio de fauna ocurrido en Australia hace unos 10 o 15 millones de años se cuenta el crecimiento en número y diversidad de los macropódidos. Algunas líneas de canguros derivaron de esta radiación relativamente reciente. Una línea de fósiles, por ejemplo, estaba constituida por canguros gigantes de cara corta y cráneo y dientes especializados en un duro ramoneo. Otra línea tenía muelas con corona relativamente alta, adaptada a la masticación de pastos. Esta última línea fue la que dio lugar a todos los grupos de grandes canguros que existen actualmente.

En resumen, los canguros están lejos de ser mamíferos primitivos y a salvo del cambio. Estos animales sólo han aparecido en escena recientemente (refiriendo este último término a la escala de los tiempos geológicos), siendo su evolución una respuesta al entorno nuevo y cambiante de una sabana semiárida en expansión. Dada la herencia genética recibida de sus antepasados insectívoros, de hace 130 millones de años, ¿cómo llegó a diferenciarse la adaptación de los dos grupos de mamíferos terios, canguros y rumiantes, siendo así que ambos estaban constituidos por animales que pacían y se enfrentaban a un mismo tipo de medio ambiente especializado?

La cría de los pequeños en una bolsa establece una diferencia obvia entre los canguros y los artiodáctilos provistos de placenta, que son su contrapartida ecológica. Para comprender por qué los canguros conservan este rasgo básico de los marsupiales hay que tener en cuenta el tipo de reproducción de estos últimos. La hembra cría a sus hijos tras un embarazo más corto que el de las hembras placentadas de tamaño equiparable. Por consiguiente, las crías nacen en estado prácticamente embrionario. No todos los marsupiales poseen marsupio o bolsa marsupial, pero todos ellos tienen hijos que permanecen firmemente pegados al pezón durante el período postnatal, espacio de tiempo que corresponde más o menos a la última etapa de desarrollo intrauterino de los placentados.

No hay unanimidad a la hora de explicar la evolución de la forma de sostén de la cría entre los marsupiales y entre los placentados. Sin embargo, como los monotremas todavía ponen huevos, se supone que dicho sostén surgió sólo después de que los primeros mamíferos se hubieran afianzado. Por ello, la discusión gira en torno a si la forma de sostén evolucionó una sola vez, antes de la separación de placentados y marsupiales, o bien lo hizo independientemente en unos



LAS VARIANTES DE LA MARCHA del canguro abarcan desde la locomoción pentápeda *(arriba)*, propia del desplazamiento lento, hasta los saltos a gran velocidad, por encima de los 40 kilómetros por hora *(abajo)*. Durante la locomoción pentápeda, la cola del canguro actúa como la tercera pata de un trípode, ayudando al animal a sostenerse sobre las extremidades anteriores mientras las posteriores se mueven hacia adelante. Durante el salto a bajas velocidades *(centro)*, la cola coadyuva al equilibrio, pero ya no funciona como soporte. El número de saltos por minuto es casi constante a velocidades comprendidas entre 15 y 35 kilómetros por hora; las velocidades superiores a la última citada se logran alargando el salto. Cuando éste adquiere gran velocidad, exige recorrer más distancia en cada brinco y saltar con mayor celeridad.

y otros. De hecho, en el equidna de hocico recto *(Tachyglossus aculeatus)*, monotrema que se alimenta de hormigas, el huevo, de textura coriácea, permanece 27 días en el útero, donde crece sensiblemente, y después es puesto en una bolsa en la que se rompe el cascarón 11 días más tarde. El componente intrauterino de este modo de reproducción aparece, pues, claramente manifiesto. No es difícil ver que el siguiente paso de la evolución estribe en la deposición de crías, en vez de huevos, dentro de la bolsa. Esta teoría es sustentada por Jason A. Lillegraven, de la Universidad de Wyoming, quien propone que todos los terios primitivos nacían indefensos y requerían un cuidado intensivo por parte de la madre. Los marsupiales han perpetuado este rasgo, pero la mayoría de placentados han prolongado el período de gestación interna, de manera que los pequeños están más desarrollados al nacer. Los ungulados han llevado a sus últimas consecuencias esta opción, pues sus crías nacen activas.

Las razones que parecen explicar el porqué de esta estrategia de nacimiento, conservada desde tiempos ancestrales y altamente desarrollada en los canguros, constituyen en parte un enigma. Entre las explicaciones propuestas citaremos la falta de una verdadera placenta y de un mecanismo que prolongue la vida activa del cuerpo lúteo del ovario. (El cuerpo lúteo segrega progesterona, la hormona responsable en buena medida de la preparación y mantenimiento del útero cuando se instaura la preñez.) Ninguna de las dos explicaciones es aceptable. En algunos marsupiales –sobre todo en el bandicut, del tamaño de un conejo– existe una verdadera placenta, a la vez que se prolonga la vida activa del cuerpo lúteo.

En otra teoría se defiende que la adaptación crucial que condujo a la prolongada gestación de los mamíferos placentados radica en el mecanismo que impide a las hembras preñadas expulsar los tejidos "extraños" del embrión. El embrión en desarrollo sería extraño a causa de su constitución doble, ya que consta de tejidos derivados no sólo de la madre, sino también del padre. Lillegraven ha señalado que, dentro del grupo de los placentados, la mayor barrera existente entre los antígenos fetales y los anticuerpos de la madre es de tipo anatómico, y consiste en una parte especial de la placenta denominada trofoblasto. En los marsupiales no se ha desarrollado esta barrera, por lo que tampoco han variado las condiciones originarias; se mantiene una cáscara de huevo inerte, naturalmente generada por la madre, hasta que la gestación está avanzada, envoltura que parece bloquear la respuesta inmunológica materna. La presencia de la cáscara limita el tamaño de la cría, pues una vez rota aquélla el embrión ha de nacer antes de que se desencadene sobre él un ataque inmunológico total.

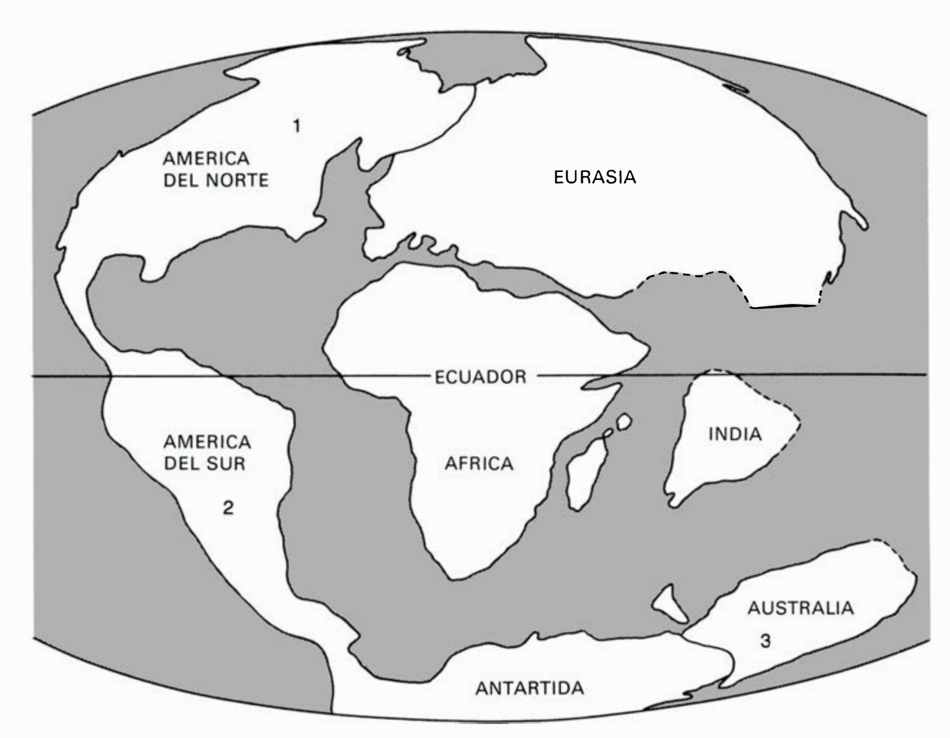
Los marsupiales han retenido la pauta reproductora básica de los terios primitivos, pero los canguros la han modificado hasta el punto de que, en algunos aspectos, incluso se puede establecer un paralelo entre dicha pauta y la de los ungulados, más perfeccionada, que implica el llevar en la matriz hijos completamente activos. En efecto, la bolsa del canguro es mucho mayor y se parece más a un saco que la bolsa más común entre los marsupiales, en tanto que el tiempo que el pequeño canguro pasa en ella se prolonga notablemente. El embarazo del canguro rojo *(Megaleia rufa)* dura solamente 33 días, dos días menos que el ciclo del estro. Su prole solamente pesa unos 0,8 g en el momento de la eclosión, pese a lo cual es mayor que la de cualquier otro marsupial, tanto en sentido absoluto como relativo. El canguro rojo todavía debe crecer mucho una vez ha trepado, completamente desasistido, desde el orificio urogenital hasta la bolsa materna. Transcurridos 235 días a partir de este acontecimiento, la cría desocupa definitivamente la bolsa.

Puede considerarse que el canguro jo-



UNA HEMBRA DE CANGURO ROJO descansa a la sombra de un arbusto de acacia en un día caluroso (temperatura del aire 46° C.) La cría que se encuentra junto a ella todavía mama, aunque hace tiempo que abandonó la bolsa. En ésta hay otra cría desarrollándose, la cual pronto se unirá a la que ya "anda" por sí misma. La postura de la madre, con la cola recogida debajo, ayuda a reducir la absorción de energía radiante.

ven ("joey", en Australia) que abandona la bolsa para siempre ha nacido por segunda vez, y cabe compararle en desarrollo al recién nacido de una hembra placentada como la cierva o la oveja. El pequeño continúa mamando durante un período equiparable al de lactancia de los ungulados. No obstante, la composición de la leche de la madre se modifica al iniciarse este período adicional de lactancia. La capacidad de las glándulas mamarias del canguro para segregar dos clases distintas de leche es una adaptación que permite a la madre amamantar a un pequeño que marcha por su propio pie, mientras simultáneamente alimenta a otro que todavía permanece en el marsupio. La cría que ya se vale por sí misma mete la cabeza en la bolsa para mamar de "su" teta, en tanto que el "joey" que sigue dentro chupa de otra. La composición y volumen de la leche segregada por una y otra mama es totalmente distinta. ¿Cómo se consigue esto, siendo el estado hormonal idéntico? Se trata de una cuestión que intriga.



EL EXODO DE LOS MARSUPIALES hacia Australia comenzó en las postrimerías del Mesozoico, cuando una parte de estos primeros mamíferos en auge abandonó América del Norte *(1)*, su patria, por América del Sur *(2)*, la cual todavía, a principios del Cenozoico, estaba unida a la Antártida y Australia. Durante la mayor parte de esta era, los dos hemisferios americanos estuvieron separados. Cuando se juntaron de nuevo, los placentados más modernos, procedentes del norte, compitieron con los marsupiales del sur, confinándolos en la reserva de Australia principalmente. Los marsupiales se extinguieron en América del Sur hacia mediados de la era cenozoica. La zarigüeya no es ningún residuo del Mesozoico, sino un emigrante desde América del Sur.

Una peculiaridad de la reproducción de muchas especies de canguros es el fenómeno conocido como "implantación retardada". Durante cierto período de descanso de la reproducción, la hembra lleva en el útero un embrión viable cuyo desarrollo se ha detenido completamente en la fase de blastocisto. Este blastocisto mide unos 0,25 mm de diámetro y se compone de unas 70 a 100 células. ¿Cuál es la función de esta adaptación? Para entenderla debemos poner atención en el ciclo reproductor del canguro.

La fecundación no altera, en general, el ciclo del estro, el cual no desaparece durante la gestación. Sólo lo interrumpe el estímulo de la lactancia, después del parto. En muchas especies, el estro y el acoplamiento se producen nuevamente poco tiempo después del nacimiento, es decir, en el momento en que se habrían producido de no haber quedado preñada la hembra. G. B. Sharman, de la Macquarie University de Australia, describió por primera vez en 1954 el destino del óvulo fecundado en una cópula inmediatamente posterior al alumbramiento.

El desarrollo del óvulo prosigue hasta la fase de blastocisto, quedando a continuación latente. En el canguro rojo, supuesto que el recién nacido sobreviva normalmente en la bolsa, el blastocisto permanecerá detenido en su crecimiento durante 200 días aproximadamente. Una vez transcurridos éstos, se reanuda el desarrollo, sobreviniendo el nacimiento 30 días después. El anterior habitante de la bolsa no tarda en ser desalojado de ella, antes de que el nuevo hijo la ocupe. Al nacimiento sigue de nuevo el estro y la cópula, de suerte que puede formarse otro blastocisto. Pero como el "joey" expulsado puede continuar mamando durante otros cuatro meses, la hembra del canguro rojo puede tener tres hijos nutriéndose de ella al mismo tiempo: un blastocisto en estado de latencia, un pequeño que se alimenta y crece en la bolsa y otro mayor, que ya "camina", pero todavía mama.

La significación adaptativa de la suspensión del desarrollo del blastocisto ha sido objeto de exhaustivos debates, pues adopta formas diferentes dentro de la familia de los macropódidos. En los canguros rojos, por ejemplo, parece que su significación iría unida a la necesidad de una rápida recuperación de la población de canguros, tras la reducción en número que provocan las sequías del árido interior de Australia. A diferencia del embrión de los placentados, conectado, para alimentarse, al torrente sanguíneo de la madre, el pequeño canguro se nutre únicamente de leche. Durante la sequía, la cría de la bolsa, e incluso la que está fuera de ella, puede perecer si disminuye el suministro de leche. Cuando ello sucede, el blastocisto reanuda su desarrollo y, tras sobrevenir el parto, la madre se aparea y engendra un nuevo blastocisto.

La cría mayor sería la primera en acusar los efectos de una sequía, y cuanto más se dilatara ésta menor sería la edad a la que moriría la cría alojada en la bolsa, hasta que al fin la hembra dejaría de criar definitivamente. Pueden verse, pues, las ventajas adaptativas de este dispositivo reproductor, gracias al cual la hembra siempre puede tener a su cargo una descendencia en desarrollo, excepto bajo condiciones severísimas. Esta adaptación posee un valor todavía mayor para una población de amplia dispersión, valor que puede apreciarse a la vista de otro hecho adicional: como ha demostrado A. E. Newsome, de la Wildlife Research Division (Commonwealth Scientific and Industrial Research Organization), los canguros se vuelven temporalmente estériles en los grandes calores y sequías.

Durante muchos años se han ido acumulando lentamente datos sobre la fisiología de la reproducción de los canguros: no obstante, sólo en fecha bastante reciente se han valorado las características de su locomoción. Los brincos de los canguros intrigaron a los primeros visitantes europeos que llegaron a Australia. Uno de los comentarios más tempranos a este respecto se encuentra en el diario de Sir Joseph Banks, quien via-

jó en el *Endeavour* con el capitán James Cook, de 1768 a 1771. Lo consignamos aquí:

"Vimos cuadrúpedos, aunque pocos, y pocos pudimos atrapar de los que vimos. El mayor de ellos era llamado *Kangaroo* por los nativos. Es distinto de cualquier animal europeo y, en verdad, de cualquier animal del que yo haya oído hablar o acerca del que haya leído, a excepción del "Gerbua" de Egipto, que no es mayor que una rata, mientras que éste mide tanto como un cordero de talla mediana; el mayor que cazamos pesaba 84 libras. Con todo, puede distinguirse fácilmente de los demás mamíferos por su singular aptitud para correr, o más bien saltar, sobre las patas traseras, llevando las delanteras dobladas sobre el pecho; sin embargo, estos animales saltan con tal rapidez por el terreno rocoso donde por lo común habitan, que varios de ellos batieron fácilmente a mi lebrel, el cual sólo pudo matar uno, bastante joven además, a pesar de haber levantado a varios."

La sorpresa de Banks ante el espectáculo de un gran animal saltador estaba completamente justificada; los pequeños saltadores como su "Gerbua" egipcio ("jerboa", es decir, el gerbo, un pequeño roedor) abundan en los desiertos de todo el mundo, pero entonces se desconocían los de gran tamaño. En realidad se habría sorprendido mucho más si hubiese sabido lo que sabemos hoy: el canguro representa probablemente el único ejemplo en la historia de la evolución de un animal de más de 3-5 kg que haya adoptado la marcha a saltos.

Durante un tiempo se creyó que los dinosaurios bípedos saltaban, pero no hay pruebas que sustenten esta hipótesis. Podemos preguntarnos por qué el salto es un sistema tan poco habitual de locomoción entre los animales grandes, y por qué los canguros lo explotaron. A medida que vamos comprendiendo cuestiones tales como la del gasto energético relativo de las diversas formas de locomoción, así como la agilidad relativa que permite cada tipo de marcha, la respuesta a las citadas preguntas va siendo evidente.

Hasta hace poco era difícil una comprensión plena de las características del salto, toda vez que no existía un marco de información general sobre el gasto energético de la locomoción dentro del cual pudiese situarse el gasto del salto. C. Richard Taylor, de la Universidad de Harvard, Knut Schmidt-Nielsen, de la Universidad de Duke, y sus colaboradores, han empezado a construir dicho marco. De sus estudios se desprende claramente que existe un gasto energético regular y susceptible de predicción para cada tipo de carrera. Este gasto depende básicamente del tamaño del animal y de la velocidad desplegada; además, pese a las diferencias que separan a bípedos de cuadrúpedos, puede preverse con bastante exactitud y se mantiene considerablemente estable, por lo menos para velocidades moderadas.

Taylor y yo teníamos intereses comunes en varios aspectos de la economía energética del canguro y de su regulación térmica, lo que nos llevó a una colaboración mutua en este campo. Nos hicimos enviar por avión cuatro canguros desde Sidney hasta Boston. Conseguimos adiestrar a dos de ellos para que saltaran en el interior de una gran rueda mientras se medía su consumo de oxígeno (y por consiguiente el gasto energético del salto). Estos estudios mostraron que el modelo energético del canguro difería acusadamente tanto de los bípedos como de los cuadrúpedos. La diferencia más significativa a este respecto es que un canguro que se desplaza a velocidad moderada lo hace con mayor economía de oxígeno que los cuadrúpedos o bípedos de tamaño similar, lo cual es una ventaja evidente para estos animales.

¿Por qué, entonces, es tan raro el salto en los animales grandes? La respuesta se relaciona probablemente con las pautas de locomoción de los canguros a velocidades bajas. A velocidades inferiores a los seis kilómetros por hora, los canguros no saltan; se mueven de forma bastante singular, sirviéndose como soporte suplementario de su pesada cola. A esta marcha se le ha dado el calificativo de pentápeda, puesto que la cola actúa como una quinta pata que ayuda a las extremidades delanteras a sostener al animal, mientras las traseras se mueven juntas hacia adelante. Contrariamente a lo que afirman opiniones muy extendidas, los canguros pueden mover sus cuartos traseros independientemente: cuando nadan –son buenos nadadores– la acción de sus extremidades posteriores es alternante.

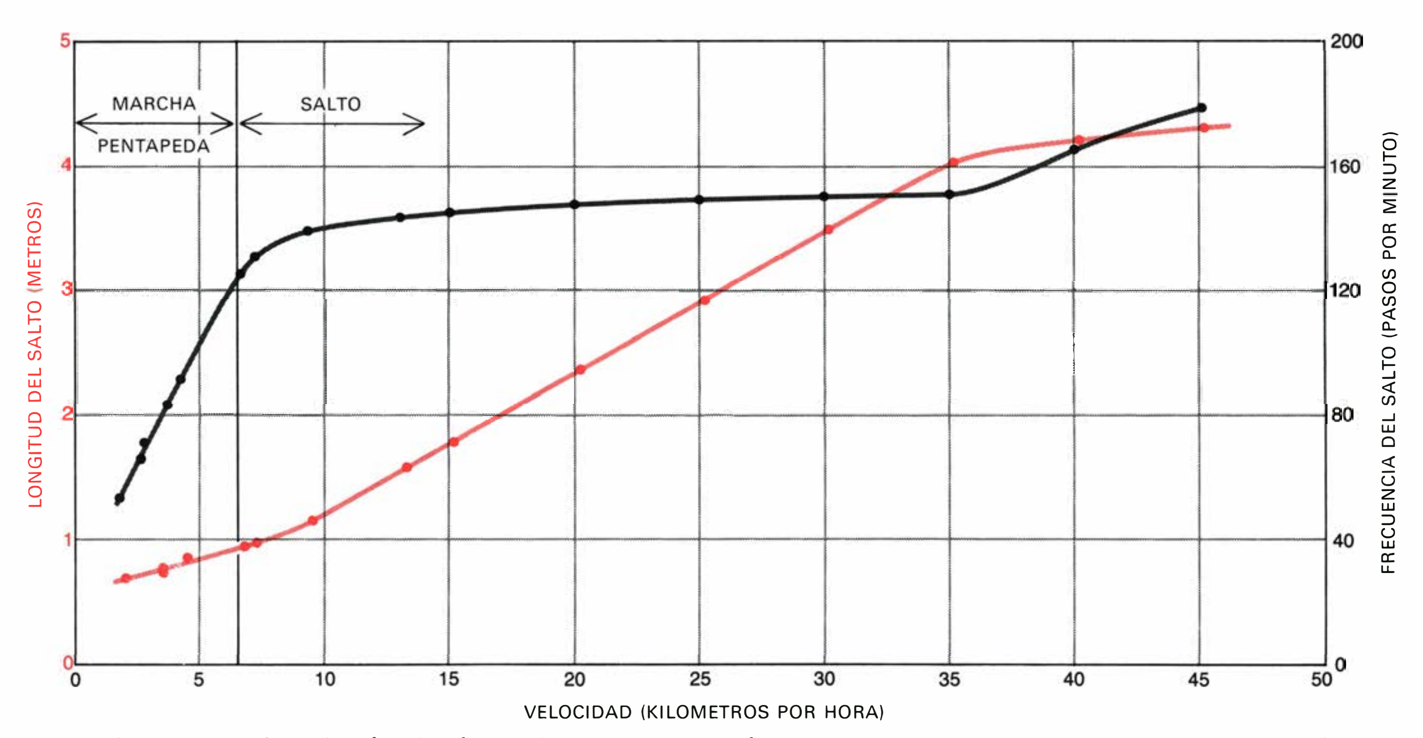
La marcha pentápeda es torpe y muy costosa desde el punto de vista energético. Pero los canguros escogieron este extraño modo de locomoción porque,

EPOCA	AUSTRALIA	AMERICA DEL SUR	AMERICA DEL NORTE	EUROPA	ASIA
RECIENTE					
PLEISTOCENO					
PLIOCENO					
MIOCENO					
OLIGOCENO					
EOCENO					
PALEOCENO					
CRETACEO SUPERIOR					
CRETACEO INFERIOR					
JURÁSICO					

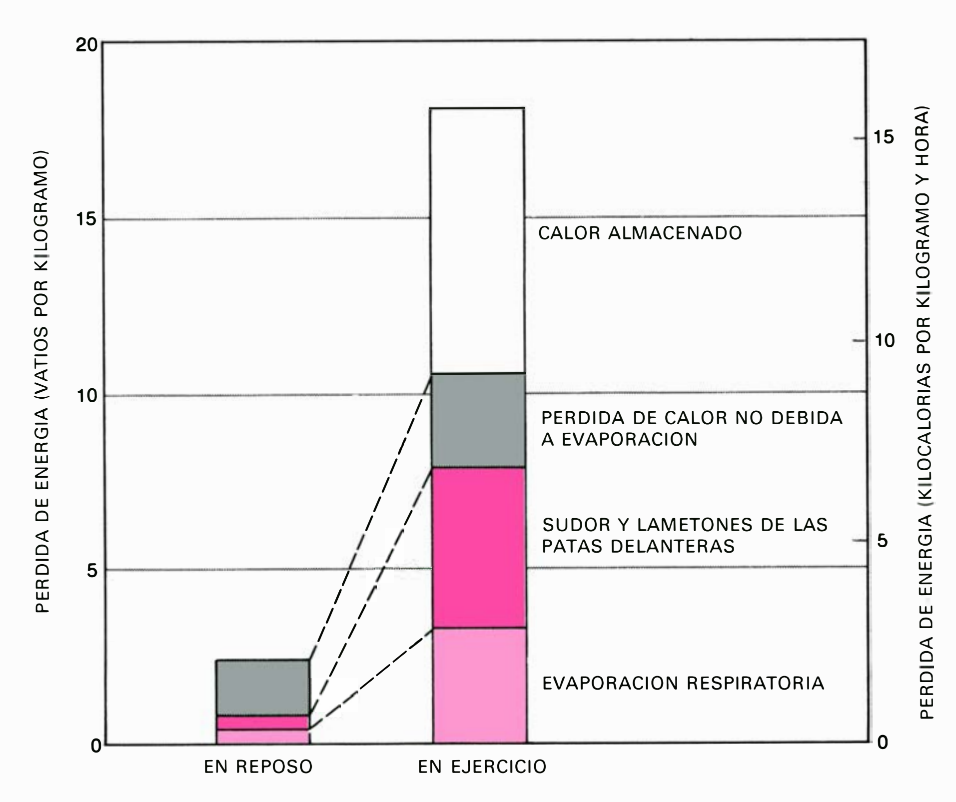
PASADO EVOLUTIVO de los marsupiales a partir de los inicios del Cretáceo, cuando las líneas de placentados y marsupiales comenzaron a divergir. Los primeros marsupiales proliferaron en América del Norte y entraron asimismo en América del Sur por esta misma época. Durante el Paleoceno, algunos marsupiales de América del Norte llegaron a Europa. En el curso del Mioceno se extinguieron ambos grupos. La mayoría de las especies de marsupiales viven hoy en Australia.

CONSUMO DE OXIGENO (LITROS POR KILOGRAMO Y HORA)
MARCHA PENTAPEDA
SALTO
LIMITE DE VELOCIDAD DE LA RUEDA
VELOCIDAD (KILOMETROS POR HORA)

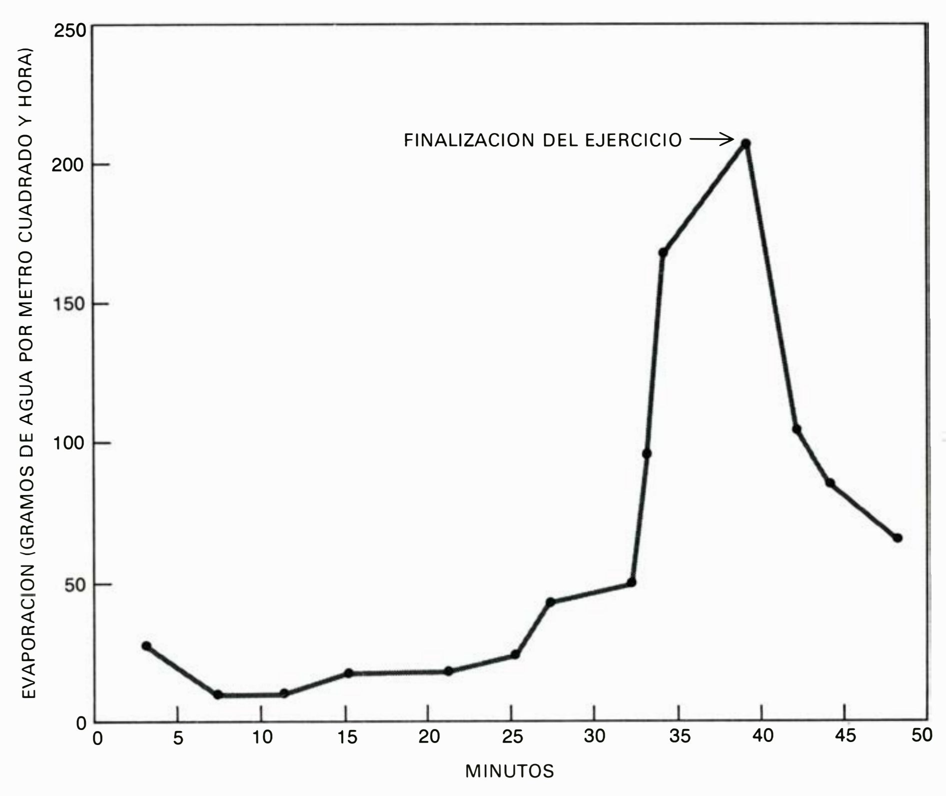
EL GASTO ENERGETICO de la locomoción, medido en función del consumo de oxígeno a diferentes velocidades, se determinó colocando a los canguros en una rueda que permitía la simulación de la marcha natural hasta los 25 kilómetros por hora. El gasto energético de la locomoción pentápeda aumentó rápidamente; ahora bien, al iniciarse el salto, el gasto se niveló e incluso decreció. La línea recta *(color)* representa la previsión del incremento de gasto debido a la velocidad en un cuadrúpedo del mismo peso. Por encima de los 17 kilómetros por hora, aproximadamente, el salto parece ser más económico que la locomoción cuadrúpeda. La prolongación de la curva *(línea discontinua)* más allá del límite de velocidad de la rueda sugiere que lo anterior puede ser cierto hasta la velocidad máxima en salto continuado (unos 45 kilómetros por hora).



LA VELOCIDAD DEL SALTO está en función de su frecuencia y longitud. La frecuencia aumenta rápidamente a medida que lo hace la velocidad en marcha pentápeda, llegando a los dos pasos por segundo antes de que el canguro comience a saltar. Sin embargo, desde los 15 a los 35 kilómetros por hora, el aumento de la velocidad del salto no va parejo con el de su frecuencia, que es trivial; en cambio, el incremento de la longitud del salto *(color)* es factor fundamental para dicho aumento de velocidad. A velocidades superiores aumenta la longitud y la frecuencia de su salto.



COMPARACION DEL EQUILIBRIO CALORICO del canguro en reposo y en pleno esfuerzo a una temperatura, moderada, de 24° C. (temperatura del aire). La pérdida de calor no debida a la evaporación basta para mantener un equilibrio estable en el canguro que reposa; las otras dos formas de pérdida de calor –por el tracto respiratorio o por enfriamiento de la piel–, combinadas, no tienen ni la mitad de importancia que aquélla. El canguro en movimiento no puede aumentar de forma significativa sus pérdidas calóricas no debidas a la evaporación; más todavía, ni siquiera un aumento pronunciado de las pérdidas causadas por evaporación basta para disipar el calor.



LA TRANSPIRACION de los canguros está correlacionada con su ejercicio físico, si bien requiere un período de calentamiento. En el gráfico se registra la aparición del primer brote de sudor en el rostro de un canguro. Este se produjo cuando el animal, forzado a moverse en la rueda en locomoción pentápeda, tuvo que avanzar a 4 kilómetros por hora. Unos 30 minutos después de haber comenzado su actividad, el canguro empezó a sudar copiosamente. Cuando finalizó el movimiento, unos 10 minutos más tarde, la transpiración disminuyó, cesando prácticamente al cabo de otros 10 minutos.

según parece, el salto a bajas velocidades aún les hubiera resultado más dispendioso. Un simple análisis físico de los diversos componentes del salto avala esta opinión. Los mismos factores explicarían también por qué el hombre varía su forma de desplazamiento, pasando a un trote lento en vez de andar más deprisa, y por qué el trepar a los árboles o correr cuesta arriba supone un mayor gasto para los animales grandes que para los pequeños. Además, la adopción de la locomoción cuadrúpeda a bajas velocidades, tal como la practican los pequeños animales saltadores, queda excluida en los canguros, debido a la burda especialización anatómica de sus miembros posteriores y a la reducción en tamaño de los anteriores. Prueba de que estas especializaciones están relacionadas con el aumento de tamaño es que algunas especies menores de los macropódidos utilizan la locomoción cuadrúpeda a bajas velocidades.

En suma, habiendo adoptado el salto sólo a velocidades medias o altas, los canguros tienen importantes problemas locomotores cuando se desplazan lentamente, pero a cambio gozan de notables ventajas sobre los animales corredores cuando utilizan la carrera a saltos. Una vez iniciada ésta, la velocidad debe elevarse considerablemente para que tenga lugar un incremento del gasto energético. En los experimentos de la rueda, cuando los canguros saltaban a más de 15 kilómetros por hora, lo hacían más económicamente que bípedos y cuadrúpedos de tamaño similar. La pauta de coste energético del salto parece explicarse por la combinación de una razón fija de salto y el almacenamiento de energía elástica en los tendones. En estos tejidos de fibra elástica la energía puede almacenarse de forma muy parecida a como se almacena en el muelle de uno de esos zancos con pie elástico que se usan para saltar. A velocidades superiores a 22 o 23 kilómetros por hora, las medidas del gasto eran impracticables en la rueda; pero, aún así, cabe hacer algunos comentarios generales que se basan en conocimientos obtenidos de animales saltadores de pequeño tamaño y de observaciones de campo sobre canguros.

Podemos anticipar la hipótesis de que los animales con terreno libre por delante se desplazan a la velocidad que les es más cómoda, es decir, a la que consume menos. Esta parece estar, para los canguros, entre los 20 y 25 kilómetros por hora; por encima de ella es probable que el gasto aumente, porque el almacenamiento de energía elástica tiene

un límite. No obstante, la frecuencia del salto permanece constante hasta alrededor de los 40 kilómetros por hora; el incremento de velocidad se logra prolongando la longitud del salto. Si se les obliga a ello, los canguros son capaces de mantener esta velocidad durante un par de kilómetros; pueden incluso superarla, aunque raramente lo hacen. Unicamente salen disparados en una serie vertiginosa de saltos cuando se da una emergencia, y aún entonces sólo recorren así unos centenares de metros. He seguido canguros que marchaban a 50 kilómetros por hora, y se ha llegado a hablar de uno que lo hacía a 65 kilómetros por hora. A velocidad máxima se percibe un marcado incremento, no sólo de la longitud del brinco, sino también de su frecuencia. Ciertamente, una elevación de la frecuencia del salto va asociada a un destacado aumento del consumo de energía, pero todavía no se sabe cómo se relaciona el gasto energético del salto con el de la locomoción cuadrúpeda, cuando ambos se efectúan a gran velocidad.

Resulta claro que la forma de locomoción del canguro tiene sus ventajas, a pesar de los inconvenientes que presenta cuando se desplaza a poca velocidad. Los canguros han experimentado una amplia radiación por toda Australia, habiendo sobrevivido no solamente a la llegada de los aborígenes y los dingos, sus perros semidomesticados, sino también a la irrupción de los europeos acompañados igualmente de sus perros. Sin embargo, el análisis biomecánico del salto indica que sus ventajas disminuyen de forma significativa al aumentar el tamaño. Acaso deba citarse aquí que, de acuerdo con nuestros conocimientos, los canguros de mayor corpulencia aparecidos en el curso de la evolución –algunos de ellos tenían más de tres metros de altura– desaparecieron paulatinamente junto con otros grandes marsupiales durante el Pleistoceno.

En todo caso, ¿por qué hizo su aparición el salto como forma de locomoción? Posiblemente la respuesta a esta pregunta estribe en el hecho de que los marsupiales poseen un rasgo conservador, acaso primitivo. Durante muchos años se pensó que las características que podían aducirse en apoyo del pretendido primitivismo de los canguros eran la escasa magnitud de su tasa metabólica, que ya se conocía, y la temperatura baja y fluctuante de los marsupiales en general. El metabolismo basal del canguro es un 70 por ciento del de los placentados más perfeccionados. Asimismo, su temperatura corporal es ligeramente inferior a la de éstos (35,5 frente a 36,5° C). Mantienen constante su temperatura.

¿Cómo conciliar una baja tasa metabólica con la energía necesaria para el salto? A este propósito deben considerarse dos aspectos de la locomoción. Uno es la velocidad máxima mantenida al escapar de un peligro. Otro es la economía del gasto en los desplazamientos a larga distancia. Las investigaciones de Taylor y sus colaboradores parecen indicar que el gasto de la locomoción cuadrúpeda es más o menos fijo. El máximo consumo de energía mantenido puede estimarse a partir del mayor volumen de oxígeno tomado, el cual, a su vez, parece estar relacionado de algún modo con el metabolismo basal. Como éste es bajo en los marsupiales, cabe dentro de lo posible que su emisión máxima de energía sea escasa, lo que se refleja, por consiguiente, en una reducida capacidad para la locomoción rápida. El salto bípedo de los canguros y otros macropódidos acaso revele un mecanismo que ayude a superar esta limitación. La economía del gasto en los desplazamientos largos podría ser importante para un animal metabólicamente limitado como el canguro, cuya zona de distribución se extiende por la región semiárida de Australia.

En el pasado se supuso que el canguro poseía una restringida capacidad de regulación de su temperatura corporal. Esta suposición se apoyaba no solamente en el hecho de que los marsupiales tienen una temperatura ligeramente inferior, sino también en observaciones realizadas sobre otro rasgo de la conducta del canguro, a saber, sus lengüetazos termorreguladores. Cuando se cazan canguros en días calurosos, aquéllos se detienen en ocasiones para lamerse las patas delanteras. Esta conducta sirvió para dar fuerza a la idea de que los canguros no podían mantener su temperatura corporal constante, dado que, por lo general, siempre se ha aceptado que el lamerse es un medio primitivo de refrigerarse. Quienes sostenían esta postura se hubieran visto en apuros para no abandonarla de haber observado a los canguros del desierto, de los que es ejemplo el rojo, en su hábitat natural.

Es difícil imaginar que un animal con deficiente regulación térmica pueda permanecer a la escasa sombra de un arbolito del desierto durante los largos días del verano. En tal situación la temperatura del aire puede sobrepasar los 45° Celsius en tanto que la temperatura ambiental llega a ser, de hecho, superior, debido a la intensidad de la radiación calórica. Se ha descubierto que estos canguros, lejos de poseer una termorregulación deficiente, hacen gala del sistema más eficaz y mejor organizado de defensa contra el exceso de calentamiento corporal que se puede encontrar en los mamíferos. Cuando les agobia el calor, los canguros jadean y sudan, aparte de lamer sus patas delanteras. La interrelación de estos sistemas de disipación del calor por evaporación sólo se advierte en el curso del ejercicio físico. Cuando el canguro descansa se libra del calor principalmente por medio del jadeo, ya que entonces no suda y sólo da lengüetazos a sus patas delanteras si el calor aprieta.

Estos lengüetazos fueron durante mucho tiempo un enigma para mis colegas y para mí mismo, pues el área lamida es habitualmente reducida. Parecía dudoso que se sacara algún provecho de la evaporación global del calor en vista de ciertas desventajas que posiblemente la acompañaban. Por encima de otras motivaciones, fue nuestra frustración la que nos hizo emprender un examen del aporte de sangre a los miembros anteriores. Gracias a la inyección de látex líquido en los vasos sanguíneos de un canguro muerto, descubrimos una red densa e intrincada de vasos superficiales en la región que los animales se lamen habitualmente. Un estudio ulterior nos reveló que cuando el canguro siente intensamente el calor aumenta en gran medida la afluencia de sangre a la mencionada región. Así pues, esta región es una zona de trasvase calórico bien diferenciada. Al extender la saliva por sus patas delanteras, los canguros están acaso sacando un partido adicional de ese fluido que les llega de más de los pulmones, lugar donde se realiza, fundamentalmente en los animales en reposo, la disipación de calor por evaporación.

La pauta de pérdida calórica varía mucho de los canguros activos a los canguros en reposo, dado que los primeros sudan. En una primera etapa de nuestra investigación propusimos que la transpiración reemplazaba al jadeo durante el ejercicio físico, por la razón de que no es posible jadear en el curso de este último. Pronto advertimos, sin embargo, que la evaporación que se produce en el sistema respiratorio es todavía más importante al moverse, porque entonces la pérdida de agua obrada a través de la superficie aumenta grandemente. De hecho, este aumento eleva la evaporación respiratoria muy por encima del nivel que el animal en reposo alcanza mediante el jadeo. Este incremento de la ventilación respiratoria representa, pro-

bablemente, una respuesta a la mayor demanda de oxígeno propia del período de ejercicio; por tanto debe venir limitada por el volumen de oxígeno incorporado. El sudor corporal parece, pues, poner en juego superficies suplementarias dedicadas a la pérdida adicional de calor por evaporación. El animal no puede lamerse mientras salta, aun cuando los canguros cesan de saltar a veces para cubrir sus miembros anteriores con saliva. Los lametones subsiguientes a la actividad física colaboran a que la temperatura normal del cuerpo se recupere rápidamente.

Hay otro aspecto de la cuestión que también separa al canguro de otros mamíferos, y es que la transpiración cesa tan pronto como el animal se detiene, ello incluso cuando su temperatura es elevada y jadea y se lame las patas delanteras. El hombre, el caballo y otros mamíferos capaces de un ejercicio continuado no sólo sudan para disipar el calor cuando están en movimiento, sino que también lo hacen en reposo. Algunos mamíferos, como el ganado vacuno en general, los carneros y otros bóvidos, jadean y sudan simultáneamente. ¿Qué gana el canguro al empezar a jadear y dejar de sudar cuando se para? Gana el agua que ahorra. Un animal que vive en un medio ambiente árido debe economizar el agua de su cuerpo, y se pierde menos agua cuando ésta se evapora por el tracto respiratorio que cuando lo hace a través de la piel. La evaporación a través de la piel causa un descenso de la temperatura superficial, lo que a su vez provoca una mayor afluencia de calor del medio ambiente al cuerpo y la utilización de una cantidad adicional de agua para eliminar este nuevo exceso calórico. El jadeo no rebaja la temperatura superficial del cuerpo.

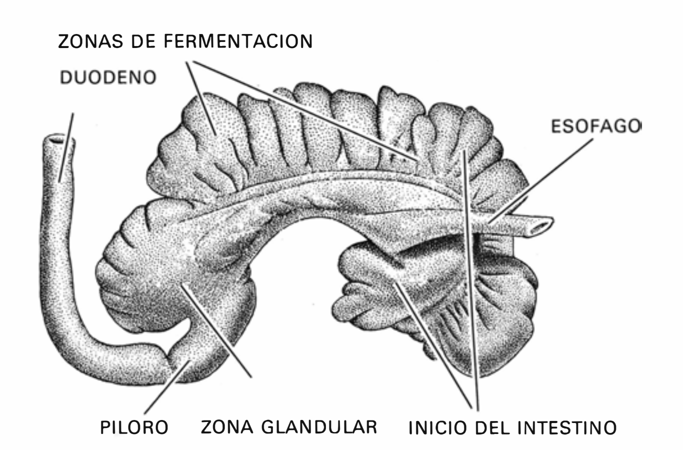
Estudiemos al canguro rojo en un día caluroso de verano. Es un ejemplo vivo de lo que puede llegar a hacer un animal para regular su temperatura corporal y conservar su agua. Cuando la temperatura sube por encima de los 45° C, el canguro rojo no yace en la sombra, sino que permanece agachado exponiendo un área superficial mínima, a fin de tomar del medio exterior la menor cantidad posible de calor. Su espesa capa pilosa le proporciona un aislante térmico ideal frente al calor ambiental. Para disminuir todavía más la radiación calórica de la superficie corporal, el canguro introduce su larga y gruesa cola entre las patas. La cola posee también una compleja red de vasos superficiales involucrados en la evaporación de calor. Según parece, este sistema deja de ser funcional cuando las condiciones favorecen el aflujo de calor sobre el cuerpo.

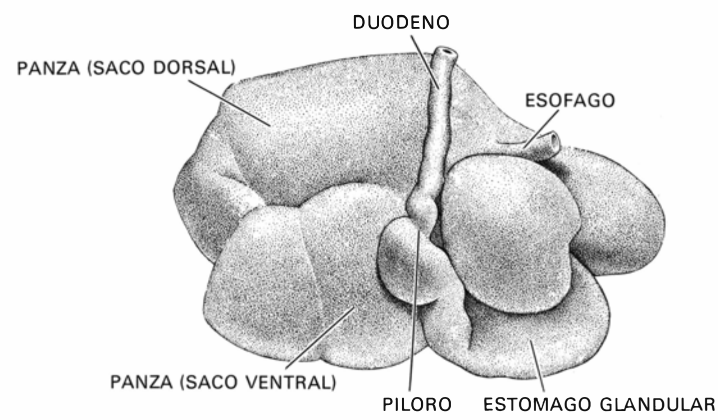
Existe una estratagema destinada a conservar el agua en condiciones de aridez que no ha sido utilizada por los canguros. Consiste en elevar la temperatura corporal durante el día. El órix, un rumiante de Africa y de Oriente Medio, posee un conjunto de vasos situados en la base del cerebro cuya disposición permite que aquél se mantenga a una temperatura tolerable, aun cuando la del resto del cuerpo ascienda hasta los 46° C. Se supone que los camellos están dotados de un dispositivo semejante. De aquí que estos dos habitantes del desierto puedan hacer acopio de calor y librarse después de él, disipándolo en el fresco del atardecer y no gastando agua en un enfriamiento por evaporación. Los canguros no disfrutan de este artificio vascular; a pesar de lo cual, su temperatura desciende por la noche si durante el día ha sufrido intensamente el calor y la deshidratación, de forma que a la mañana siguiente pueden reanudar su vida con dos o cuatro grados menos en el cuerpo.

Como consecuencia de esta combinación de adaptaciones al ahorro de agua, las necesidades mínimas de los canguros a este respecto son similares a las de los ungulados de los desiertos africanos. Los estudios de campo indican que los canguros de zonas áridas necesitan tan sólo un cuarto del agua que exige el metabolismo de ovejas y cabras silvestres. El órix puede prescindir completamente del agua de beber; el líquido que necesita lo toma de las plantas que come y de la oxidación del producto alimenticio. El canguro, al igual que el camello, debe beber ocasionalmente cuando el calor impone condiciones muy severas.

La visión de la fisiología de los marsupiales ha estado teñida constantemente por la idea de primitivismo. Considérese el hecho de que el riñón de un canguro que habita en zonas áridas tiene una capacidad de concentración superior a la de otros mamíferos placentados de hábitat semejantes. En realidad esto no debería sorprender a nadie: las características fundamentales del riñón de los mamíferos parecen ser muy antiguas. Hasta los monotremas poseen un buen riñón de tipo mamífero, y bien pudiera ser que los terápsidos, reptiles antepasados de los mamíferos, lo tuvieran también. En otros aspectos de la economía del agua los canguros resisten igualmente bien la comparación con los placentados. Canguros y camellos, por ejemplo, presentan una respuesta semejante a la deshidratación: cuando pierden agua, modifican el volumen del sistema vascular a fin de no exponer el corazón a riesgo alguno.

La alimentación y la digestión del canguro entrañan adaptaciones cuya comparación evolutiva con las corres-





LA DIGESTION DE LA CELULOSA queda facilitada por la acción de los enzimas segregados por las bacterias estomacales. Esta capacidad evolucionó independientemente en los canguros (marsupiales) y en los rumiantes (placentados). A la izquierda, el estómago de un canguro rojo y, a la derecha, el de un buey. La molécula de celulosa, presente en la dieta herbívora de ambos, se rompe, en buena parte, por la fermentación.

pondientes a los rumiantes rinde interesantes resultados. Ambas líneas evolutivas han llegado a una solución similar del problema que supone nutrirse de los tejidos de las plantas. Se ha sugerido una y otra vez que la evolución de los mamíferos estuvo estrechamente correlacionada con la de las plantas con flores. A principios del Cenozoico estas plantas eran en gran parte de tipo forestal. En el Eoceno las hierbas y los pastos habían empezado a extenderse, y en el Mioceno, cuyo comienzo se sitúa hace unos 25 millones de años, grandes áreas de sabana herbácea cubrían ya las regiones más secas de la mayoría de interiores continentales. La adaptación a pacer de los canguros parece ir asociada a esta ampliación de las áreas de pasto en Australia, de modo paralelo a la expansión de los artiodáctilos placentados en el resto del mundo.

Ningún mamífero puede sintetizar los enzimas que rompen la molécula de celulosa, el principal hidrato de carbono de la hierba. La utilización de esta fuente de alimento exige la evolución de una simbiosis entre el mamífero y las bacterias y protozoos que disponen de tales enzimas. Los microorganismos se alojan en diversos sectores dilatados del intestino de los mamíferos. En los caballos y conejos viven en zonas terminales, como el ciego. En los rumiantes viven en la primera parte del intestino, o en zonas ampliadas y especializadas del esófago y el estómago. Se deja, pues, que los materiales vegetales fermenten en estas cámaras antes de alcanzar el estómago, lo que ofrece marcadas ventajas; de este modo, tanto los productos de la fermentación como los microorganismos pueden ser digeridos y absorbidos a lo largo de todo el intestino delgado. La mayor eficacia de este sistema da tal vez cuenta, por lo menos en parte, del siguiente hecho: los herbívoros que predominaban en el Eoceno eran animales, como los caballos, en los que la fermentación tenía lugar al final del intestino, mientras que aquéllos fueron desplazados en el Mioceno por una masiva radiación de mamíferos, como los ungulados, cuya zona de fermentación estaba en el inicio del intestino.

Los canguros pertenecen a este último grupo, de modo que comparten con el ganado vacuno, las ovejas, las cabras, los ciervos y los camellos las ventajas de dicho tipo de digestión. Entre esas ventajas se encuentra una menor demanda de proteínas. La urea, formada por la rotura de las moléculas de proteína, es devuelta con la saliva a las cámaras anteriores del estómago de los rumiantes, donde vuelve a ser usada para sintetizar proteínas nutritivas por los microorganismos que allí viven, en vez de verterse toda ella en la orina y, por tanto, perderse. Este doble circuito de reutilización permite a canguros y rumiantes, no sólo extraer el alimento de las secas hierbas, pobres en proteínas, de las zonas áridas, sino también guardar la reserva corporal de agua que, de lo contrario, se necesitaría para la excreción de la urea. La combinación de este eficaz sistema de digestión con la baja tasa metabólica de los marsupiales otorga a los canguros un formidable repertorio de características adaptativas, las cuales lo ponen en condiciones de explotar con fortuna el medio ambiente de la sabana.

Así pues, la visión que tenemos de los canguros es la de un grupo de mamíferos que no son en absoluto primitivos, sino que se han adaptado y han tenido una reciente radiación en respuesta a un medio ambiente nuevo y en proceso de cambio. Es cierto que estos animales poseen características que pueden considerarse primitivas, pero también lo es que han hecho suyos diversos medios de sortear las limitaciones de tales rasgos. Además, en los canguros, al igual que en otros marsupiales, se han desarrollado capacidades que sobrepujan las de los placentados de tipo comparable. Los marsupiales y los placentados son parientes próximos; su evolución por separado ha durado unos 130 millones de años y arroja luz sobre cómo pueden aplicarse soluciones evolutivas diferentes a problemas similares. Se dio un salto desde los diminutos mamíferos del Cretáceo superior, parecidos a musarañas, hasta los mamíferos con pezuñas del Viejo Mundo y de las dos Américas, por un lado, y los canguros australianos, por otro. Los testimonios históricos parecen señalar que hay más canguros actualmente en Australia que antes de llegar los europeos. De este modo, es posible que, lejos de hallarse en vías de extinción, los canguros –por lo menos los grandes canguros de la zona semiárida– estén en una fase de prosperidad, a pesar incluso de la competencia con el hombre y con los rebaños de rumiantes que éste ha introducido. Ello se debe, probablemente, a la ampliación del hábitat árido que le es adecuado. Esta ampliación, sin embargo, ha reducido simultáneamente el hábitat propio de los representantes más pequeños del grupo, los wallabies y las ratas canguro. Son estos marsupiales los que han desaparecido en buena parte del interior de Australia.

Los vacúolos gasíferos de las cianofíceas

Las células de muchas especies acuáticas de estos organismos que parecen bacterias contienen miles de diminutas estructuras llenas de gas. Dichas estructuras regulan la flotabilidad del alga

A. E. Walsby

Se sabe que al menos 50 de entre los muchos cientos de especies de plantas primitivas agrupadas bajo la denominación de cianofíceas (algas verdeazuladas) contienen unas estructuras brillantes de forma irregular denominadas vacúolos gasíferos. Estas estructuras se encuentran solamente en microorganismos procariotas, es decir, bacterias y cianofíceas (también conocidas como cianobacterias). Además, se localizan únicamente en los procariotas de vida libre que habitan en estanques, lagos y el océano; no suelen darse en los procariotas que viven en corrientes, en el suelo y en todos los demás hábitats terrestres. Esta distribución nos indica cuál sea la función de los vacúolos gasíferos, a saber: proporcionar flotabilidad y regularla. Para organismos como las cianofíceas, que viven en los cinco metros superiores, más o menos, de la columna de agua y dependen de la luz solar para la fotosíntesis, la capacidad de regular la flotabilidad es una propiedad importante. Los detalles de la estructura de los vacúolos gasíferos y del elegante modo como estos orgánulos microscópicos son regulados por las cianofíceas han sido aclarados recientemente, después de casi una centuria de investigación.

La investigación de la función de los vacúolos gasíferos comenzó en 1895, cuando el microbiólogo alemán Hans Klebahn informó acerca de un notable experimento que había realizado con una suspensión de cianofíceas obtenidas de la superficie de un lago en el que había tenido lugar un intenso desarrollo de la población algal. Había llenado una botella de vidrio recio, hasta el borde, con la suspensión: había colocado un tapón de modo que no quedara ningún espacio con aire ni burbujas visibles en la botella y, después, había golpeado el tapón con un martillo. Ocurrieron tres cosas: el aspecto de las algas cambió inmediatamente de un verde lechoso a un verde translúcido y oscuro; se acumularon burbujas de gas bajo el tapón; y, por último, las algas se hundieron gradualmente hacia el fondo de la botella. (Las algas no tratadas de esta manera flotaban hacia la superficie.).

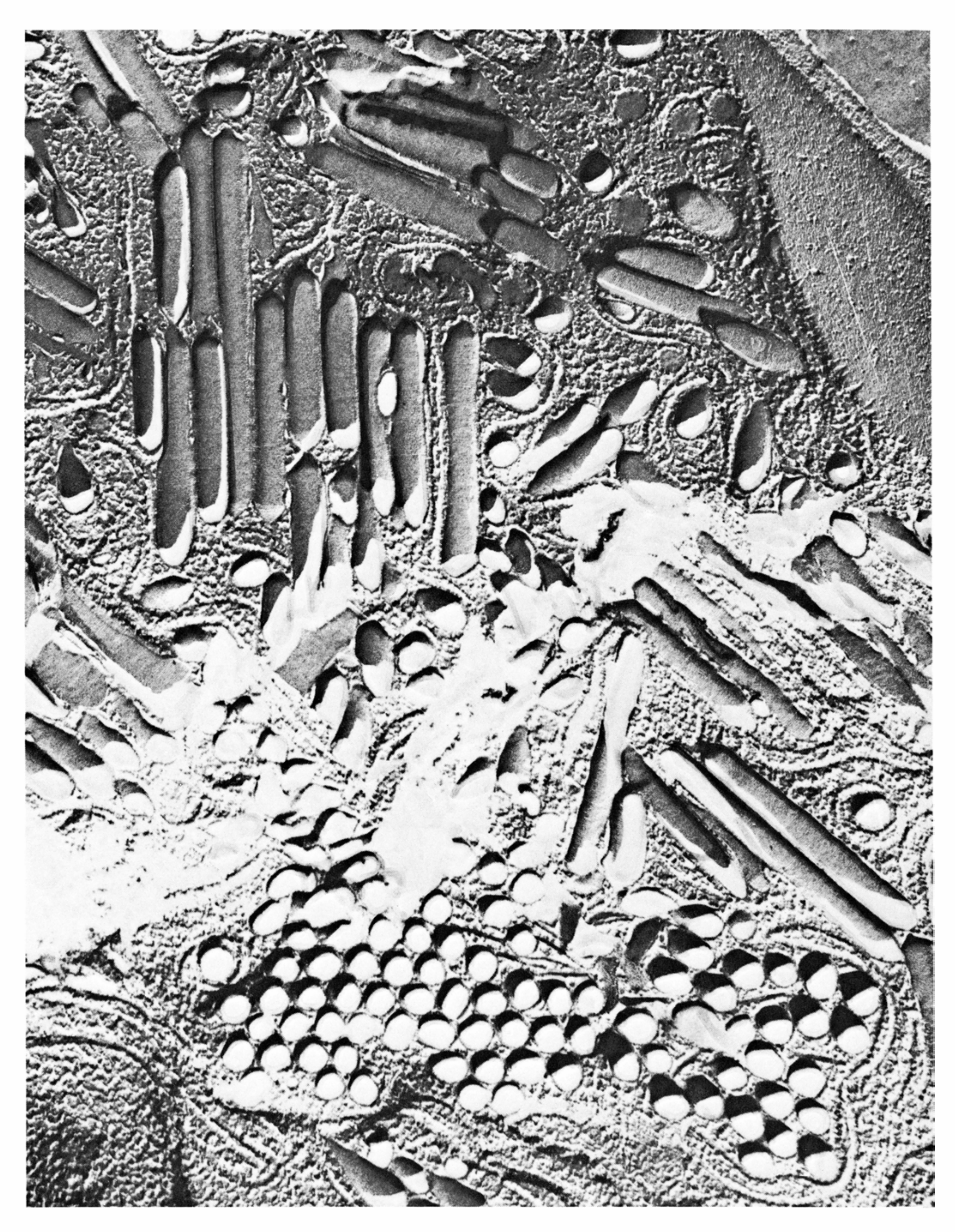
El examen de la suspensión en el microscopio mostró que la presión aplicada por el martillazo había ocasionado la desaparición de numerosos gránulos refringentes y brillantes de las células algales. Klebahn concluyó que los gránulos debían haber estado llenos del gas desprendido de la suspensión, y dedujo que la destrucción de aquéllos explicaba el cambio de turbidez de la suspensión y la pérdida de flotabilidad de las algas. Denominó vacúolos gasíferos a las estructuras.

Durante los 30 años siguientes, Klebahn trató de extraer, por diversos medios, el gas de los vacúolos y determinar su composición. En este trabajo, supuso que una membrana impermeable rodeaba el vacúolo e impedía el escape del gas, hipótesis que le permitía explicar el hecho de que no desaparecieran los vacúolos mediante exposición al vacío. (Si una solución que contiene burbujas normales de gas se coloca en un vacío, los gases se difunden fuera de las mismas, que luego desaparecen cuando la solución vuelve a la presión atmosférica.) El concepto erróneo del vacúolo como estructura impermeable a los gases duró más de 50 años. Al parecer, se centró la atención en la naturaleza, modo de formación y función del gas de los vacúolos más que en la estructura del vacúolo mismo, con lo que, en el transcurso de este período, se registraron escasos avances en la comprensión de dicha estructura.

En 1965, C.C. Bowen y T. E. Jensen, de la Universidad estatal de Iowa, publicaron en *Science* un extracto de su estudio de la ultraestructura de los vacúolos de las cianofíceas. Hallaron que cada célula algal contenía un gran número de diminutas estructuras cilíndricas huecas con extremos cónicos. Bowen y Jensen probaron que estas estructuras, que llamaron vesículas gasíferas, eran los bloques del armazón de los vacúolos gasíferos al demostrar que las vesículas desaparecían bajo presión, dejando láminas membranosas en su lugar. Evidentemente, las estructuras huecas se habían aplastado.

Me enteré de este descubrimiento cuando trabajaba en el departamento de G. E. Fogg, en el Westfield College de la Universidad de Londres, donde varios de nosotros investigábamos la fisiología de las cianofíceas. La lectura crítica del artículo me incitó a reexaminar el problema del gas en los vacúolos. Fogg me proporcionó sus traducciones de los primeros escritos alemanes sobre la materia, que él había revisado 25 años antes. Allí leí la conclusión de Klebahn de que el gas era principalmente nitrógeno. Con su análisis químico no había podido desechar la posibilidad de que el gas inerte argón estuviera también presente. Era importante verificar este punto, porque el argón no es un producto del metabolismo; si aparecía, sólo podría haber entrado en el vacúolo por difusión hacia su interior desde la solución circundante.

En aquel tiempo, estaba estudiando la fijación de nitrógeno y empleaba un espectrómetro de masas para analizar las proporciones de los diversos isótopos del nitrógeno. Creí que con este instrumento sería fácil distinguir el nitrógeno del argón en el gas de los vacúolos. Para ob-



VESICULAS GASIFERAS de la cianofícea *Anabaena flos-aquae*, aumentadas 89.000 veces en esta fotografía tomada con el microscopio electrónico. El material se congeló y, después, se fraccionó; en el proceso, unas vesículas se rompieron en el sentido de su eje mayor y, otras, perpendicularmente. Las estructuras en forma de cigarro de la parte superior son vesículas fracturadas longitudinalmente; los círculos pequeños, próximos a la parte inferior, son vesículas fracturadas transversalmente. Un grupo tal de vesículas constituye un vacúolo gasífero.

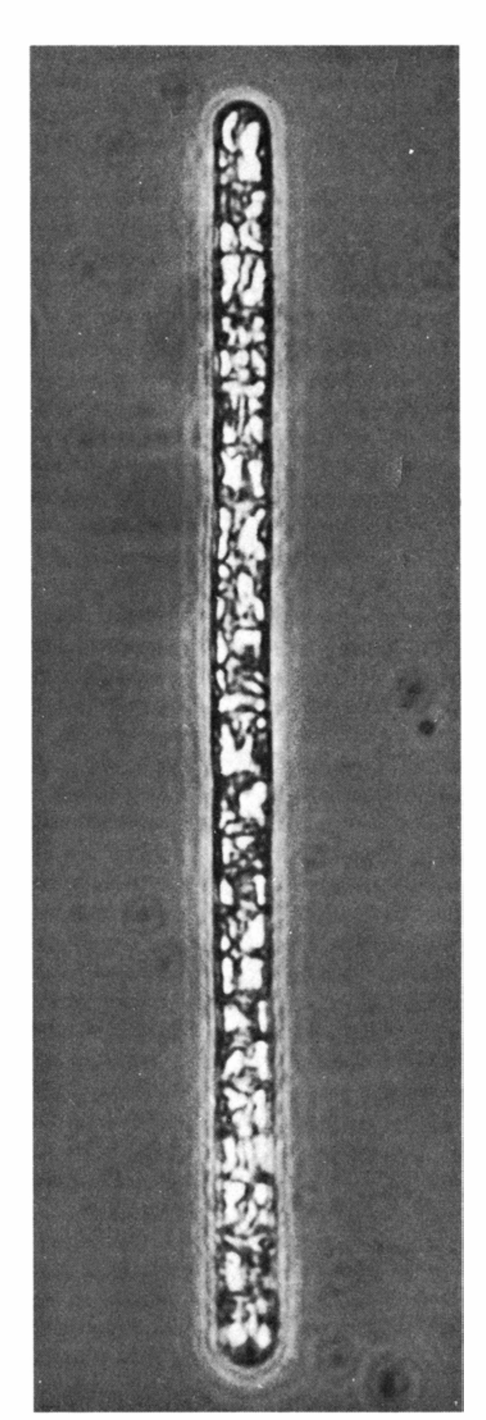
tener el gas modifiqué una de las técnicas de Klebahn; puse primero una suspensión concentrada de algas en un vacío a fin de eliminar el gas disuelto y calenté luego la suspensión en un recipiente cerrado herméticamente para romper los vacúolos. Para consternación mía, sólo pude detectar trazas o nada de gas.

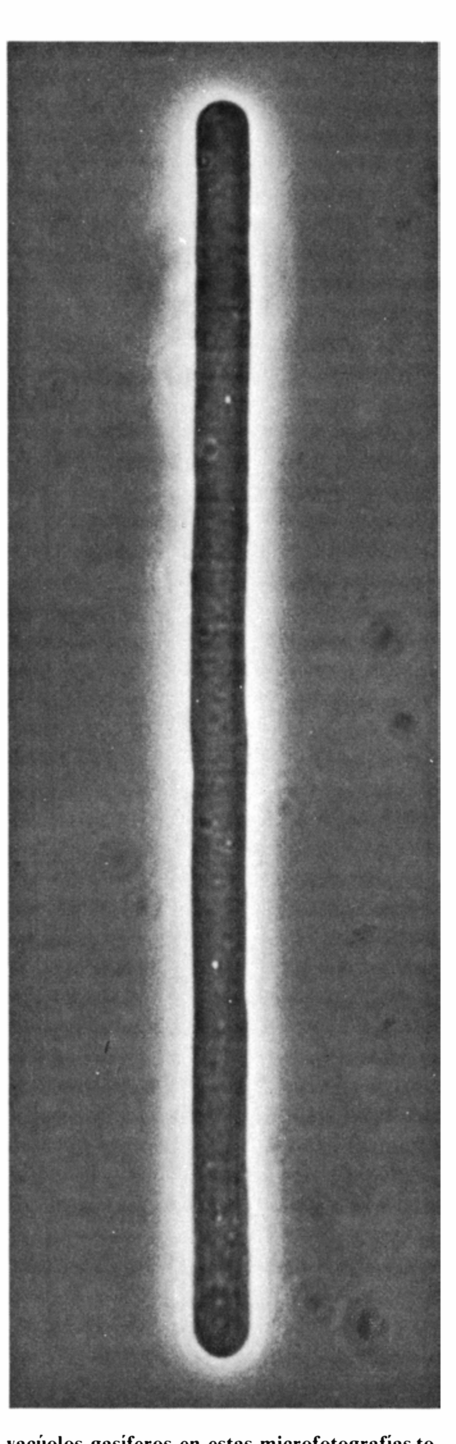
En un momento dado, observé que el gas había salido de los vacúolos al exponer las algas al vacío. En otras palabras, los vacúolos gasíferos (más exactamente, sus vesículas constituyentes) no eran impermeables, sino libremente permeables a los gases.

Ironías de la ciencia, no llegué a este descubrimiento con un instrumento de la tecnología moderna, sino con un aparato del tiempo de Klebahn, el respirómetro de Warburg. Con él medí el volumen de gas que podía obtenerse de los vacúolos en una suspensión concentrada de algas. Hallé que el volumen era siempre directamente proporcional a la presión del gas disuelto en el agua en que las células estaban suspendidas. (La presión de un gas disuelto, según se sabe, es igual a la presión del gas sobre el líquido cuando el sistema se ha equilibrado mediante una agitación enérgica.) Modificando adecuadamente el aparato de Warburg, pude seguir el intercambio de gas a través de las membranas de las vesículas gasíferas. Si se aumentaba la presión del gas sobre el líquido, se disolvía más gas en el agua y se difundía a través de la membrana porosa de la vesícula gasífera hacia el espacio situado dentro de ella; si se disminuía la presión, el flujo se invertía. Así fue como llegué a la demostración de la permeabilidad de la membrana a una serie de gases.

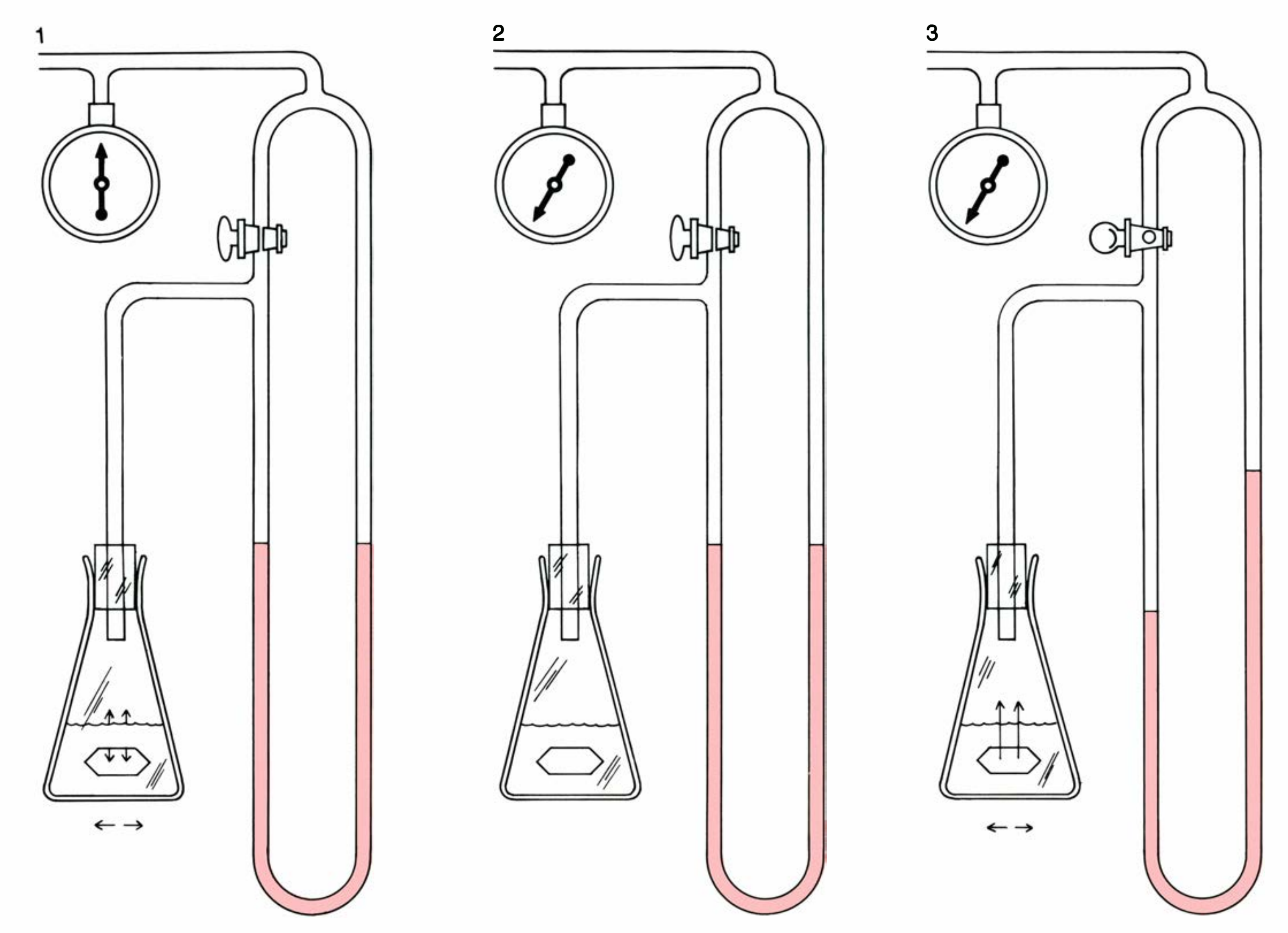
Este descubrimiento mostró que carecía de sentido tratar de extraer el gas de los vacúolos, ya que tendría siempre la misma presión y composición que los gases disueltos en la solución circundante. El hallazgo planteaba también una serie de problemas nuevos. ¿Qué era lo que impedía que las vesículas gasíferas se aplastaran a una presión moderada? ¿Cómo se formaba el espacio hueco? ¿Por qué no se llenaba de agua? Ya que la vesícula no podía almacenar un gas determinado, ¿en qué estribaba su función? Para hallar las respuestas tuve que ahondar en muchas ramas de la biología, desde la bioquímica hasta la ecología.

Klebahn había deducido correctamente que el gas de los vacúolos tenía que estar rodeado por una membrana rígida, pues podía hallarse a una presión inferior a la presión hidrostática en la célula algal. (Los azúcares, aminoácidos, iones y demás elementos disueltos en el agua del interior de la célula generan una presión osmótica, de modo que el agua del exterior tiende a difundirse hacia el interior de la célula y ésta se hincha. El hinchamiento es contrarrestado por la pared celular, generándose así una presión de reacción: la presión hidrostática o de turgencia.) Además, con el microscopio óptico, Klebahn no pudo detectar cambios en el tamaño de un vacúolo gasífero al aplicar una presión moderada. De ello no se deduce necesariamente que el volumen de las vesículas constituyentes no se altere, pero he podido confirmar que muestran muy poca deformación elástica al cambiar la presión. Por una parte, los filamentos algales (tiras de células algales) o colonias que flotan con vacúolos de gas no suben más rápidamente cuando la sus-





FILAMENTO ALGAL, que aparece con y sin vacúolos gasíferos en estas microfotografías tomadas con contraste de fases. En la fotografía de la izquierda, los vacúolos se ven como gránulos brillantes, cada uno de los cuales está constituido por varios cientos de vesículas cilíndricas huecas. Aunque las vesículas son demasiado pequeñas para que puedan verse al microscopio óptico, se hacen visibles porque los espacios llenos de gas que contienen dispersan la luz. La fotografía de la derecha muestra el mismo filamento después de haber sido sometido a una presión de 12 atmósferas. La presión hizo que las vesículas se aplastaran, por lo que ya no son visibles. El filamento (cianofícea de la especie *Oscillatoria agardhii*) está aumentado aproximadamente 1400 veces.



EL RESPIROMETRO DE WARBURG fue el aparato con el que el autor determinó que la membrana de la vesícula no era impermeable, sino permeable a los gases. El instrumento consta de un pequeño frasco conectado a un manómetro de vidrio (un tubo en forma de *U* que contiene agua coloreada) por medio de una juntura impermeable a los gases. Vacúolos gasíferos aislados de una cianofícea se suspendieron en agua dentro del frasco a la presión atmosférica *(1)*. El frasco se agitó *(flechas inferiores)* para equilibrar el gas en el agua y en los vacúolos *(flechas superiores)*. Se cerró el aparato y se practicó rápidamente el vacío *(2)* para reducir la pérdida de gas de la suspensión por difusión. Finalmente, la llave del manómetro se cerró *(3)*, con lo que el gas desprendido de la suspensión situada en el frasco empujó hacia abajo el fluido de la parte izquierda del tubo del manómetro. El cambio de presión se calcula a partir de la diferencia de altura del fluido en los dos lados. Un centímetro equivale a una presión de 0,001 atmósferas. Si las vesículas hubieran resultado impermeables no se hubiera detectado gas.

pensión se pone en un vacío. Por otra parte, la cantidad de luz dispersada por una suspensión de vesículas intactas y aisladas no muestra prácticamente cambios reversibles al aumentar o disminuir la presión ejercida sobre ellas.

Una característica importante, por lo que respecta a la rigidez de la vesícula gasífera, debe ser su forma cilíndrica, similar a la forma de los submarinos o a la de las cubas cilíndricas diseñadas para almacenar gases comerciales. Estas estructuras están dibujadas a propósito para resistir la presión. En el Westfield Collage, con Henry Eichelberger, observé que las vesículas gasíferas mantienen su forma incluso al ser aisladas de las células, lo que probaba que las vesículas no basan su soporte en el protoplasma que normalmente las rodea.

Un ejemplo sacado de la vida cotidiana puede aclarar la naturaleza de la vesícula. No es como un balón, cuyo tamaño está determinado por el exceso de presión del gas que contiene. Antes bien, se parece a una cañería porosa y rígida que tenga cerrados los extremos. La presión del gas del interior está determinada por las presiones parciales de los gases disueltos en el agua circundante y es independiente de la presión hidrostática que actúa en el exterior de la estructura. La diferencia entre estas dos presiones viene soportada por la pared de la vesícula, que es relativamente fuerte, aunque frágil, por lo que se aplasta bajo una presión suficiente. La estructura permanece estable en un vacío, aun cuando el gas se difunda hacia fuera de la misma.

Decidí ver cuánta presión se necesitaba para aplastar las vesículas gasíferas, medición que sorprendentemente no se había realizado hasta entonces. Apliqué una presión, que iba incrementándose a razón de 0,5 atmósferas, a una suspensión del alga *Anabaena flos-aquae*, provista de vacúolos gasíferos, y medí la turbidez de la suspensión después de cada paso. Observé que, para cada incremento por encima de aproximadamente una atmósfera, había un descenso de la turbidez que podía atribuirse al aplastamiento de una parte de las refringentes vesículas gasíferas de cada célula. Cuando la presión era ya de unas 4,5 atmósferas, no quedaban vesículas intactas.

Una vesícula podría aplastarse instantáneamente a cierta presión, a su presión crítica, pero a presiones inferiores permanecería intacta indefinidamente. Cada célula contenía varios miles de vesículas gasíferas. La presión bajo la que se aplastaban se distribuía normalmente en torno a un valor medio de 2,5 atmósferas, aproximadamente.

Sabía desde el principio que las presiones críticas que yo estaba midiendo subestimarían la verdadera resistencia de las vesículas, porque éstas ya soportaban la presión de turgencia dentro de las células algales. Pude eliminar la presión de turgencia y medir la distribución de la verdadera presión crítica suspendiendo las células en una solución concentrada de sacarosa, que invertía la relación osmótica normal y provocaba el flujo de agua al exterior de las células. Resultó que oscilaba desde 4,5 a 7,5 atmósferas, aproximadamente.

A consecuencia de la realización de tales mediciones, conseguí las primeras estimaciones dignas de confianza de la presión de turgencia en células procariotas. Los cálculos implicaban solamente sustraer, de la presión verdadera, la presión crítica aparente de las vesículas gasíferas. Con este método examiné una serie de organismos provistos de vesículas gasíferas y hallé una interesante correlación: casi en cada caso, el rango de presiones críticas era aproximadamente una atmósfera superior a la presión de turgencia de la célula. La diferencia proporcionaba un margen de seguridad adecuado para impedir el aplastamiento de las vesículas bajo la presión hidrostática a la que el organismo vive. Recientemente encontré una excepción interesante en la cianofícea marina *Trichodesmium*. Tiene vesículas gasíferas extremadamente fuertes, capaces de soportar la presión que el alga experimenta cuando es arrastrada a considerables profundidades del océano.

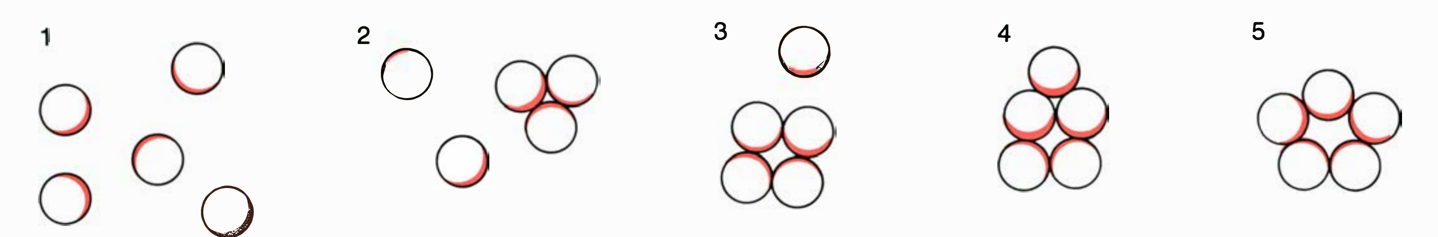
Mientras medía la presión de turgencia en *Anabaena flos-aquae* obtuve un premio inesperado que me abrió un nuevo campo de investigación sobre los vacúolos gasíferos. Observé que en las soluciones concentradas de sacarosa las células algales revientan, al parecer porque la disminución de su volumen dobla las paredes celulares. Vi que de cada célula salían empujándose miríadas de vesículas, cada una como un pequeño punto luminoso en el microscopio de contraste de fases. Esta observación constituyó la base para el aislamiento de vesículas de gas intactas, a partir de las células. (Los métodos tradicionales de fraccionamiento celular generan fuerzas que causan el aplastamiento de las frágiles vesículas.)

Aprendí de Pal Falkenberg, que trabajaba sobre halobacterias (bacterias que prosperan en ambientes salinos), en el laboratorio de Helge Larsen, en la Universidad Técnica de Noruega, que las vesículas gasíferas se podían separar mediante centrifugación de los demás componentes de la célula. Al centrifugar las vesículas, flotarían hacia la superficie, de donde podrían ser extraídas. El problema principal era que la centrifugación generaba una presión que aplastaría las vesículas. Entonces se hundirían. Sin embargo, como podía determinar la presión crítica mínima de las vesículas gasíferas en mi preparación, podía asegurarme de que esta presión no fuera excedida en la centrífuga, con lo que obtendría un buen rendimiento.

Con Barbara Buckland (estudiante graduada) obtuve preparaciones de vesículas gasíferas altamente purificadas, centrifugando las vesículas y filtrándolas después a través de un filtro de membrana. Nos dedicamos luego a analizar su composición química, con resultados sorprendentes. El grupo de Larsen, en Noruega, y el de Walter Stoeckenius, en la facultad de Medicina de la Universidad de California, en San Francisco, trabajando con vesículas de halobacterias parcialmente purificadas, había informado de que las proteínas eran un componente importante de la vesícula. Nosotros sólo pudimos hallar proteína. Estaba claro que la membrana de las vesículas, si podía ser denominada así, era distinta de las membranas típicas de la célula, que contienen a menudo hasta un 50 por ciento en peso de lípidos o grasas.

Mientras estábamos afanados en ello, Daniel D. Jones y Michael Jost, de la Universidad estatal de Michigan, aislaban vesículas gasíferas de la cianofícea *Microcystis aeruginosa*. Poco después de haber escrito nosotros un breve informe de nuestro trabajo en *Nature*, publicaron ellos un relato de sus análisis más completos, que no sólo confirmaban que las proteínas eran el único constituyente de la vesícula, sino que también mostraban que sólo estaba presente un tipo de molécula proteica. Resultaba ser relativamente pequeña (14.000 daltons). A través de cuidadosas mediciones con el microscopio electrónico, estos investigadores demostraron que la pared o membrana de las vesículas tenía el grosor de una sola molécula proteica: las moléculas se hallaban ordenadas en filas a lo largo de las fajas que formaban la estructura cilíndrica. De estos estudios surgió un modelo de la vesícula gasífera como una estructura biológica de simplicidad ini-

RELACIONES DE PRESION de una vesícula en una célula algal. La presión total en el exterior de la vesícula es de 6,5 atmósferas: una atmósfera de la presión normal en la superficie del agua, 1,5 atmósferas de la presión del agua a una profundidad de 15 metros y cuatro atmósferas de la presión de turgencia, generada por la resistencia de la pared celular a la tendencia de la célula a hincharse al incorporar agua. La presión del gas en la vesícula es de aproximadamente una atmósfera, por lo que la pared de la vesícula debe soportar una diferencia de presión de unas 5,5 atmósferas. Para cada vesícula existe un punto (la presión crítica) en el que la diferencia excede a la resistencia de la pared y ocasiona el aplastamiento de aquélla. En la vesícula intacta la presión crítica es de seis atmósferas; en la vesícula inferior esta presión era de cinco atmósferas.



LA FORMACION DE LA VESICULA comienza cuando moléculas de proteína (esferas) se reúnen en la célula algal. La proteína tiene una superficie *(color)* hidrófoba (que repele el agua) y otra hidrófila (que atrae el agua). Originariamente, las moléculas estaban suspendidas en la célula *(1)*. Al comenzar la agregación *(2 y 3)*, las superficies hidrófobas se orientan hacia el interior, de modo que el espacio hueco que se forma *(4 y 5)* tiende a eliminar el agua. Una vez formado éste, el gas se difunde hacia el interior de la vesícula a partir del agua circundante.

gualada. Los estudios de difracción de rayos X realizados por Allen Blaurock, en sus trabajos con W. Wober en la facultad de Medicina de la Universidad de California, en San Francisco, y conmigo, en la Universidad de Londres, han revelado determinados detalles de la ordenación de las moléculas de proteína en la membrana.

Estos conocimientos han permitido responder, al menos en parte, a las preguntas que formulé más arriba sobre las propiedades, formación y función de las vesículas gasíferas. Séanos permitido considerar ahora por qué el agua no entra dentro de la vesícula. No es necesario postular que la pared proteica sea impermeable a las moléculas de agua; todo lo que se necesita es que la superficie interna de la vesícula sea hidrófoba. Esta propiedad impediría que se formaran gotitas de agua dentro de la estructura y generaría una tensión superficial suficiente para evitar que se escurriera agua líquida a través de cualquier poro que pudiera haber en la membrana.

¿Puede una proteína proporcionar una tal superficie hidrófoba? Sí, si expone apropiadamente en este lado sus aminoácidos alifáticos (aminoácidos con una cadena lateral grasa). Los estudios de difracción de rayos X realizados por Blaurock pusieron de manifiesto que esta ordenación puede existir en la proteína de la vesícula gasífera. Los recientes trabajos de David Worcester, del Atomic Energy Research Establishment en Harwell, Inglaterra, han confirmado esta ordenación. Por sugerencia de Blaurock, dejó que preparaciones de mis vesículas de *Anabaena* aplastadas, que habían sido desecadas en una superficie plana, absorbieran agua pesada: agua en la que el hidrógeno no es el isótopo común del elemento sino el isótopo pesado deuterio. Las preparaciones fueron seguidamente expuestas a un haz de neutrones, que podían distinguir entre el deuterio del agua pesada y el hidrógeno ordinario en el material circundante. De las imágenes de dispersión de neutrones podía deducirse que el agua pesada no penetraba entre las dos mitades opuestas de la membrana aplastada, es decir, las superficies dirigidas hacia el espacio lleno de gas en la estructura intacta. El agua ordinaria se comportaría de la misma manera.

¿Qué hay sobre la superficie externa de la vesícula gasífera, la dirigida hacia el agua? Resulta más probable que sea hidrófila, con lo que se minimizaría la presión producida por la tensión superficial, que tendería a aplastar la estructura. Se han registrado varias observaciones que indican que la superficie externa de la vesícula gasífera es mojable. Por ejemplo, si se agitan vesículas gasíferas aisladas con aceite de oliva y agua (cosa que hice por sugerencia de Fogg), quedan totalmente en el agua tras la partición.

¿Cómo se forma el espacio hueco dentro de la vesícula gasífera? Cuando me di cuenta de que la permeabilidad de la membrana excluía la posibilidad de que estuviera hinchada de gas, propuse que la vesícula tenía que ser una estructura capaz de autoconstruirse. Su formación podría visualizarse del modo siguiente: el punto de partida sería un grupo pequeño y compacto de moléculas de la proteína de la membrana de la vesícula: la adición de nuevas moléculas apartaría a las primeras para crear un espacio hueco, los gases podrían difundirse hasta este espacio desde la solución circundante; la energía necesaria para la formación de la estructura viene de la disminución de la energía libre de las moléculas de proteína, y la atracción mutua de las moléculas impide que caigan hacia el interior del espacio hueco. J. Robert Waaland y Daniel Branton, de la Universidad de California en Berkeley, obtuvieron pruebas de formación *de novo* a través de las fases siguientes en el crecimiento de vesículas gasíferas inducido en la cianofícea *Nostoc muscorum*. La formación comenzó con una pequeña estructura bicónica que, al alcanzar un diámetro adecuado, creció por medio de la creación y elongación de la porción central cilíndrica.

La proteína de las vesículas gasíferas parece tener un papel puramente estructural, a saber: crear y mantener un espacio hueco. Este espacio realiza la función atribuida a los vacúolos gasíferos. Los acontecimientos que acompañan a la destrucción de los vacúolos en el experimento del martillo, el tapón y la botella han dado pie para pensar en varias funciones posibles.

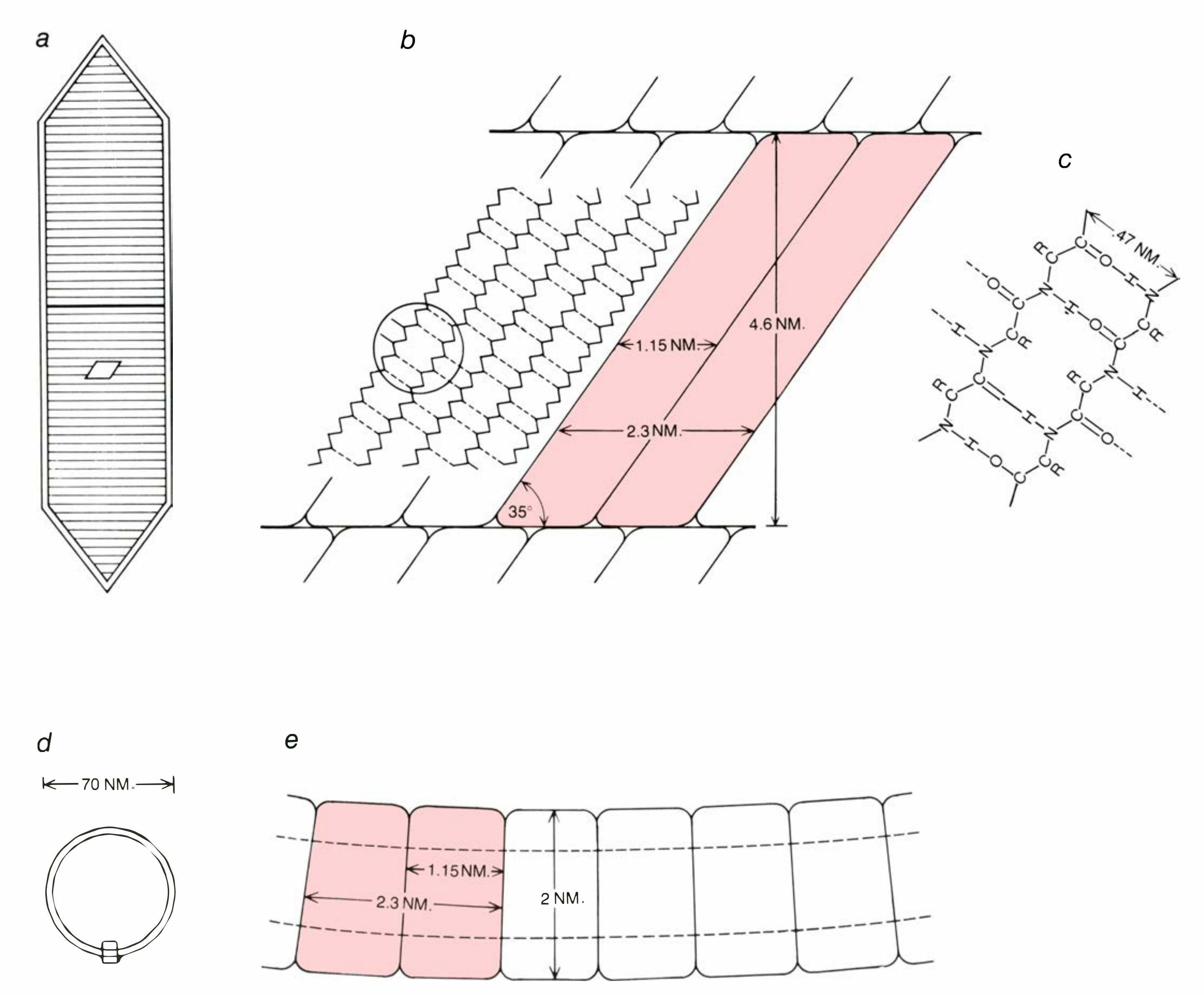
La formación de burbujas de gas en el experimento sugería que la función de los vacúolos era el almacenamiento de gas, pero esta posibilidad debe descartarse a la luz del descubrimiento de que los vacúolos son demasiado permeables para almacenar gas alguno. El cambio de aspecto de las algas en el experimento sugería una función de protección de la luz, en la que los vacúolos defenderían a las substancias sensibles a la luz de las células algales. Esta función podría tener un valor considerable para las cianofíceas que flotan en la superficie del agua bajo el sol veraniego. Sin embargo, la cuestión de la protección de la luz no está clara. Si las vesículas gasíferas son eficientes para sombrear, ¿por qué no se encuentran en las algas que habitan en la superficie de las rocas y en el suelo? Mi opinión es que si los vacúolos tienen un papel en la protección de la luz, es secundario a su función principal, que consiste en proporcionar flotabilidad.

El experimento del martillo, el tapón y la botella demuestra claramente que los vacúolos gasíferos proporcionan flotabilidad. Y, si nos fijamos en su densidad, observaremos que lo realizan con gran eficacia. La densidad de una vesícula gasífera es de 0,1 gramos por centímetro cúbico, aproximadamente. Como el agua llena el espacio que queda entre las numerosas vesículas que constituyen un vacúolo, la densidad del vacúolo completo es de unos 0,2 gramos por centímetro cúbico, mucho menor todavía que la del líquido o sólido más ligero hallado en células vivas.

La distribución de los vacúolos gasíferos entre los organismos procariotas ofrece una fuerte prueba circunstancial de que la flotabilidad que proporcionan es importante. Los vacúolos se hallan casi exclusivamente en microorganismos acuáticos planctónicos (suspendidos en el agua). Este punto lo subraya el hecho de que, siendo prácticamente desconocidas las bacterias con vacúolos gasíferos entre los miles de cepas no acuáticas que se mantienen en laboratorios de todo el mundo, pude hallar cerca de 30 especies diferentes de bacterias provistas de vacúolos en un solo estanque de Minnesota.

Los hábitats más comunes para las bacterias dotadas de vacúolos gasíferos son los lagos relativamente estables que llegan a presentar estratificación térmica, como ha confirmado recientemente Alison Clarck (estudiante graduada que trabaja conmigo) en un extenso reconocimiento de masas de agua en Gran Bretaña. Durante el verano, se desarrollan poblaciones de estas bacterias desde el fango del fondo y se suspenden en el hipolimnion (la capa de agua más fría del fondo de un lago). La población de ciertas bacterias presentan máximos a una determinada profundidad, donde probablemente las condiciones son las más favorables para su crecimiento. Por ejemplo, las bacterias fotosintéticas requieren condiciones de luz tenue y ausencia de oxígeno, que en un lago estratificado se pueden hallar justo por debajo de la termoclina (la zona de transición entre la capa caliente superficial y el agua fría situada bajo ella). Por otra parte, las cianofíceas planctónicas pueden formar poblaciones estables hacia la parte inferior del epilimnion (la capa superficial caliente).

Estos fenómenos sugieren que los vacúolos gasíferos pueden estar haciendo algo más que simplemente disminuir la densidad de estos organismos; también pueden estar implicados en la regulación activa de la flotabilidad, de modo que las células puedan suspender-



LA ESTRUCTURA PROPUESTA para una vesícula gasífera en una cianofícea se ilustra de acuerdo con análisis químicos y estudios de difracción de rayos X. En *a* la vesícula aparece en sección longitudinal. Las paredes tienen un grosor de dos nanómetros y cada faja tiene 4,6 nanómetros de anchura. El área de fajas delimitada por la caja se muestra aumentada en *b*, donde se ve que cada faja está compuesta de células unitarias de 1,5 nanómetros de anchura con una inclinación de 35 grados. Cada molécula proteica incorpora probablemente dos de estas células unitarias *(color)*. La ultraestructura de parte de la molécula se muestra a la izquierda; cada cadena polipeptídica está arrollada en ocho largos paralelos dispuestos en dos capas, sólo una de las cuales aparece aquí. El área rodeada por el círculo se muestra con mayor detalle en *c*, donde se han ilustrado los enlaces químicos de la molécula. La vesícula completa aparece en sección transversal en *d*, y el área de pared delimitada por la caja se reproduce en *e* a escala mayor. Las líneas discontinuas representan las capas de la cadena polipeptídica. El área coloreada muestra un modelo plausible de molécula proteica; mide 2,3 nanómetros de anchura y consta de dos células unitarias.

se en determinadas ubicaciones de una columna vertical de agua. Las algas planctónicas que mantenía en cultivo para estudiar los vacúolos gasíferos me dieron la primera pista sobre cómo podría conseguirse la regulación de la flotabilidad.

Me di cuenta de que si un alga crecía a una baja intensidad luminosa, flotaba invariablemente, mientras que bajo una alta intensidad de luz se hundía. Advertí que, si el alga se comportaba del mismo modo en un lago, quedaría estratificada a una profundidad con una intensidad de luz intermedia. Al continuar investigando, hallé que el alga perdía su flotabilidad muy deprisa (en una hora más o menos) si se la transfería de una baja a una alta intensidad luminosa.

La fase siguiente fue fortuita. Cuando trataba de establecer el rango de la presión de turgencia en esta alga por medio de mi método de los vacúolos gasíferos, observé que la turgencia era mayor en células algales dejadas a una alta iluminación que en las dejadas a baja iluminación. Se me ocurrió que, en las algas transferidas a mayores intensidades luminosas, la presión de turgencia podría aumentar lo suficiente como para aplastar las vesículas más débiles, con lo que se explicaría la pérdida de flotabilidad. Posteriormente, M. T. Dinsdale y el autor, en el Westfield College, obtuvieron pruebas directas del sospechado aumento de turgencia y demostraron que dependía de la fotosíntesis. Se había comenzado a desentrañar la cadena de hechos implicados en esta respuesta.

Durante los últimos veranos he estudiado las cianofíceas en los lagos para ver si regulan su flotabilidad de este modo. Los resultados indican que lo hacen así. Por ejemplo, con Andrew Klemer, de la Universidad de Minnesota, examiné un lago en el norte de Minnesota que contenía una población permanente de la especie de cianofícea *Oscillatoria agardhii;* durante el verano mostraba su máximo a la profundidad de 5 metros, cerca del fondo de la zona eufótica (la zona donde hay suficiente luz para permitir el crecimiento de las plantas verdes). Las mediciones realizadas con un nefelómetro de presión (un aparato portátil que diseñé para medir los vacúolos gasíferos y la presión a la que se aplastan) mostraron que, a aquella profundidad, las algas tenían justamente los vacúolos gasíferos suficientes para quedar neutralmente suspendidas. ¿Era esto accidental, o se debía a que el alga regulaba su flotabilidad?

Ideamos un experimento sencillo para demostrar que el alga regulaba realmente su flotabilidad. Tomamos muestras del agua del lago a una profundidad de cinco metros, donde la población algal era unas 1000 veces más concentrada que en la superficie, pusimos el agua en botellas de vidrio que tapamos, atamos a una cuerda y suspendimos a varias profundidades desde una plataforma flotante. Al día siguiente, recuperamos las botellas y encontramos que en la botella de cinco metros las algas habían permanecido homogéneamente suspendidas. En las botellas colocadas cerca de la superficie, las algas se habían hundido y, en las botellas situadas por debajo de cinco metros, habían tendido a subir. Las diferencias persistían cuando se dejaban las botellas a temperatura constante, lo que mostraba que habían tenido lugar cambios permanentes en la densidad de las algas. Posteriores observaciones sugirieron que el cambio en la densidad resultaba de un cambio en la dotación de vacúolos gasíferos, que a su vez ocurría como respuesta al decrecimiento de la intensidad luminosa con la profundidad. La dotación de vacúolos gasíferos del alga parecía regulada por el mecanismo de la presión de turgencia que he descrito.

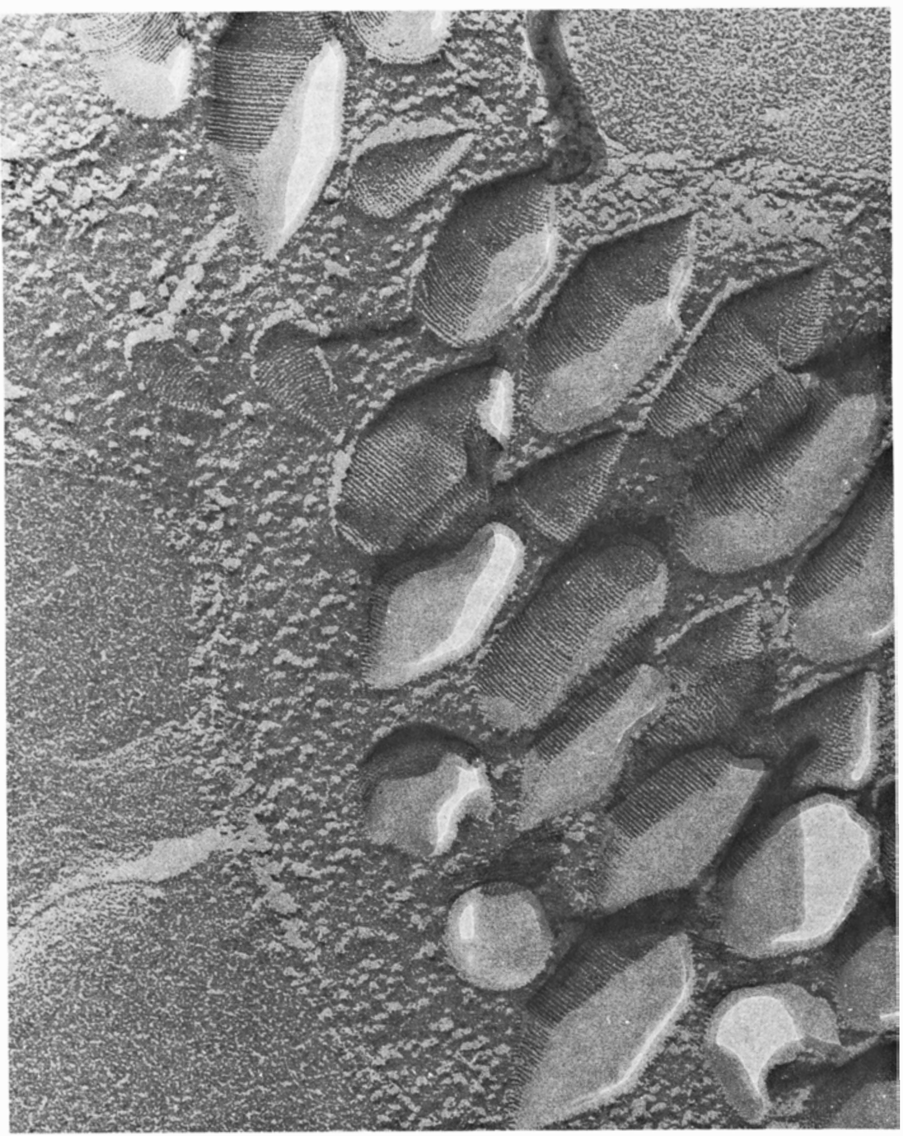
El mismo mecanismo está también implicado en las migraciones verticales diurnas del alga *Aphanizomenon flos-aquae* en el lago Mendota, en Wisconsin. Investigué estas migraciones con Allan Konopka, del laboratorio de T. D. Brock, de la Universidad de Wisconsin. El lago Mendota es demasiado turbulento como para quedar estratificado de un modo permanente, y el alga suele hallarse homogéneamente suspendida en el epilimnion. Sin embargo, durante los períodos de calma, el alga responde a los cambios de intensidad luminosa subiendo hacia arriba durante la noche y hundiéndose durante el día. *Aphanizomenon* puede hacer esto porque forma grandes colonias que (por razones hidromecánicas) se mueven mucho más deprisa a través del agua que los diminutos filamentos de *Oscillatoria.* Otra cianofícea colonial, *Microcystis aeruginosa,* realiza migraciones similares, pero quizá mediante un mecanismo diferente. La cantidad de vacúolos gasíferos puede decrecer a altas intensidades luminosas, porque se forma nuevo material celular a un ritmo mayor que el registrado en las nuevas vesículas.

Estos estudios han permitido comprender la estratificación algal y la migración vertical. Sin embargo, un tercer fenómeno parece oponerse a estas observaciones. Es la formación de los denominados "blooms", densas acumulacio-

nes de cianofíceas que aparecen en la superficie de muchos lagos durante el verano. Colin Reynolds, de la Freshwater Biological Association, en Gran Bretaña, ha demostrado que estas acumulaciones no se forman por el crecimiento de algas en la superficie; el crecimiento ocurre antes, mientras el alga está suspendida en el epilimnion. La acumulación ocurre durante períodos de calma, cuando las algas, con flotabilidad positiva a causa del exceso de vacúolos gasíferos, emergen hacia la superficie. La formación de una de estas acumulaciones es paradójica, en el sentido de que a menudo las algas mueren en la superficie por agotamiento de nutrientes minerales y exposición a la intensa luz solar.

Reynolds y el autor, después de revisar material publicado sobre estas acumulaciones algales, llegaron a la conclusión de que su formación se debía a la incapacidad del alga de reducir su vacuolación gasífera. Esta incapacidad reflejaba la imposibilidad de mantener altas tasas de fotosíntesis en la superficie, a pesar del elevado nivel de iluminación existente, ya que ambos métodos de regulación de los vacúolos gasíferos (presión de turgencia y crecimiento celular) dependen de la actividad fotosintética. Dedujeron así la existencia de varias causas para este fallo de la fotosíntesis.

Una de ellas vino sugerida por experimentos que Michael Booker y el autor hicieron en un gran depósito de laboratorio que, con iluminación por encima y refrigeración en el fondo, pretendía simular un lago estratificado. Observaron que, cuando una cianofícea planctónica crece en el depósito bajo condiciones en que los nutrientes son limitantes, se estratifica a un nivel en que la intensidad luminosa es baja. Si el agua se enriquece con nutrientes, fosfato especialmente, la población algal aumenta y, por medio de su actividad fotosintética, consume el anhídrido carbónico disuelto en el agua. Como la fotosíntesis requiere anhídrido carbónico, queda limitada su continuación, y el alga resta intensamente vacuolada y sube a la superficie, acumulándose allí. No se sabe si una explicación parecida es válida para las aguas naturales, pero es significativo que los lagos donde las cianofíceas se estratifican mejor son oligotróficos (es decir, tienen bajos niveles de nutrientes tales como el fosfato), mientras que las acumulaciones masivas tienden a formarse en aguas eutróficas (con altos niveles de nutrientes).



VESICULAS GASIFERAS de una bacteria, *Prostheconicrobium pneumaticum*, **aumentadas 139.000 veces en esta microfotografía electrónica hecha por Branton mediante la técnica de congelación y fraccionamiento. Algunas especies acuáticas de bacterias tienen vesículas muy similares a las halladas en ciertas cianofíceas. Se aprecia la forma de la vesícula (cilíndrica con extremos puntiagudos) y las fajas, de sólo 120 nanómetros de anchura, que forman las moléculas proteicas de la pared vesicular. Las vesículas tienen 120 nanómetros de anchura y hasta 300 nanómetros de longitud; las de las cianofíceas poseen una forma similar, si bien resultan mucho más estrechas (70 nanómetros) y más largas (llegan hasta los 1000 nanómetros).**

Las grandes proliferaciones de cianofíceas causan considerables daños. Interfieren los suministros de agua al obstruir los filtros y desprender substancias que producen sabores desagradables. Pueden afectar adversamente a las pesquerías de agua dulce. Algunas cianofíceas formadoras de estas acumulaciones sintetizan poderosas neurotoxinas que pueden causar la muerte del ganado y de animales salvajes. Además, los "blooms" echan a perder las aguas utilizadas para fines recreativos.

En muchos casos, estos problemas se deben a la acción humana, ya que las condiciones eutróficas que soportan estas acumulaciones de algas se deben al enriquecimiento de lagos y embalses en nutrientes procedentes de aguas residuales y de fertilizantes aplicados en exceso a terrenos agrícolas. Es posible que un mejor conocimiento del vacúolo gasífero y de su papel en la biología de las cianofíceas planctónicas conduzca a hallar algún medio de controlar los "blooms". D. C. Menday y A. A. Buck, de la Universidad de Gales, han demostrado que mi sugerencia, un tanto humorística, de que estas acumulaciones podrían ser contenidas aplastando los vacúolos gasíferos con explosivos, es económicamente viable, aunque, desde luego, es inaceptable por otras razones. Mi propia opinión es que el problema de las acumulaciones masivas de algas sólo puede resolverse invirtiendo el proceso de eutrofización en las aguas en que aparezcan.

Perspectivas en química macromolecular

La hipótesis macromolecular ha permitido tanto el desarrollo de la gran industria de los plásticos como penetrar en los secretos de la vida. Hoy es necesario ampliar nuestros conceptos y métodos para abrir nuevos caminos

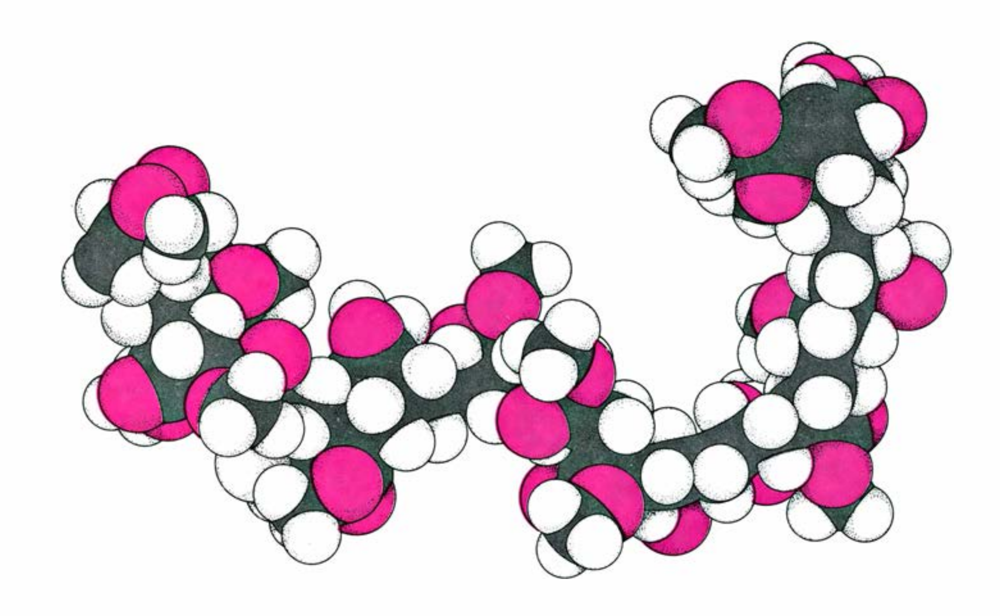
Juan Antonio Subirana

Nuestra vida cotidiana se desenvuelve en un mundo saturado de materiales plásticos. Los envoltorios, los electrodomésticos, los asientos de los coches, los indicadores de las autopistas están formados en su mayor parte por estos materiales. Su resistencia a la destrucción ambiental hace que nos los encontremos en rincones perdidos de la montaña o en playas desiertas, como residuo de una civilización que no sabe disponer adecuadamente de sus desperdicios. Esta resistencia a la destrucción se debe fundamentalmente a que los materiales plásticos están formados por moléculas gigantes, es decir, por macromoléculas. Lo que quizá no nos hemos detenido a pensar es que la resistencia de nuestra piel, de nuestras uñas, de nuestros músculos obedece a la misma causa, es decir, a la naturaleza macromolecular de nuestros tejidos. Químicamente, los componentes de nuestro cuerpo y los materiales plásticos guardan un estrecho parentesco, con algunas diferencias notables, evidentemente. En estas páginas vamos a comentar las perspectivas actuales en este campo, tanto en el desarrollo de los materiales plásticos, como en la mejor comprensión de los fenómenos biológicos cuyo vehículo son las macromoléculas.

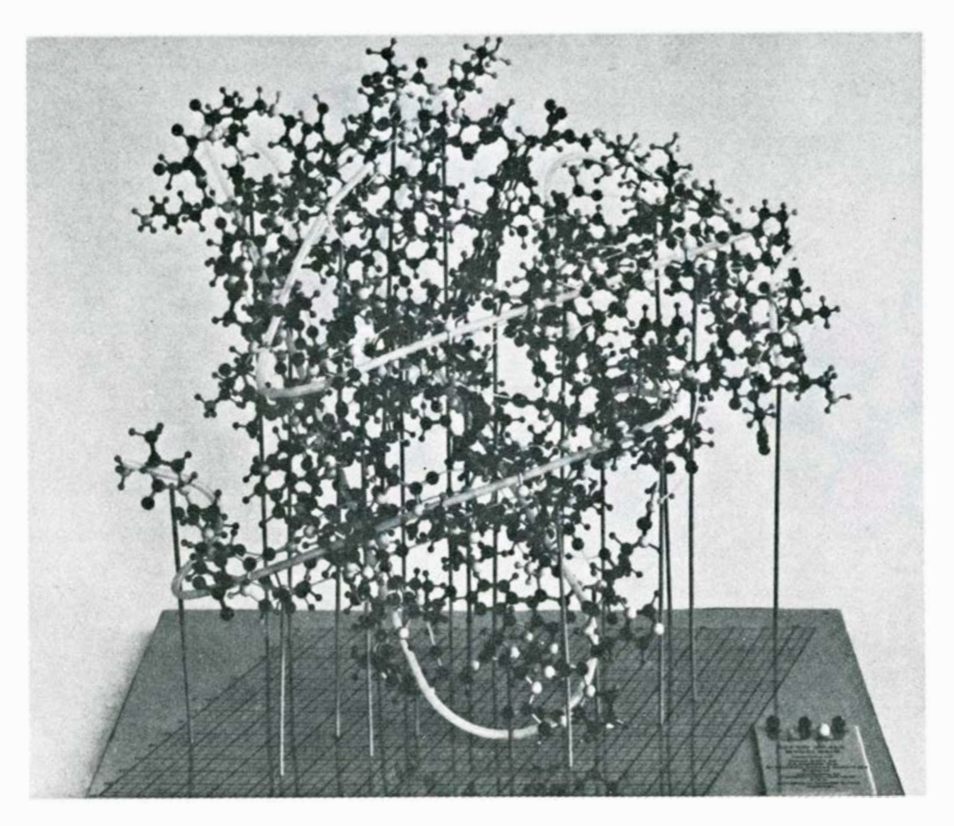
Principios estructurales

Un hecho a primera vista sorprendente es que el concepto de macromolécula no se introdujo hasta alrededor de 1920 y no se aceptó plenamente hasta 1930. Al haberse establecido ya en el siglo pasado la teoría del enlace covalente, según la cual los átomos se unen entre sí mediante electrones apareados, no parecía que pudiera existir ningún obstáculo conceptual a la idea de que podían existir macromoléculas. Sin embargo, algunos de los dogmas imperantes en la ciencia de aquellos años dificultaron que se aceptara esta hipótesis. Uno de estos dogmas postulaba en que cualquier substancia pura ha de cristalizar y tener un punto de fusión definido. Esta era la filosofía de los químicos orgánicos; por lo que todas las substancias macromoleculares, que en general no pueden cristalizar, eran desechadas por ellos como substancias impuras, de poco interés. Esta metodología dificultó que se aceptara la posibilidad de la existencia de macromoléculas. Pero este hecho, y volveré al final sobre ello, demuestra que en un determinado estado de conocimientos hay una serie de dogmas que se aceptan sin demasiadas pruebas, dogmas que actualmente también existen, y que pueden obstaculizar el progreso hacia nuevos avances científicos.

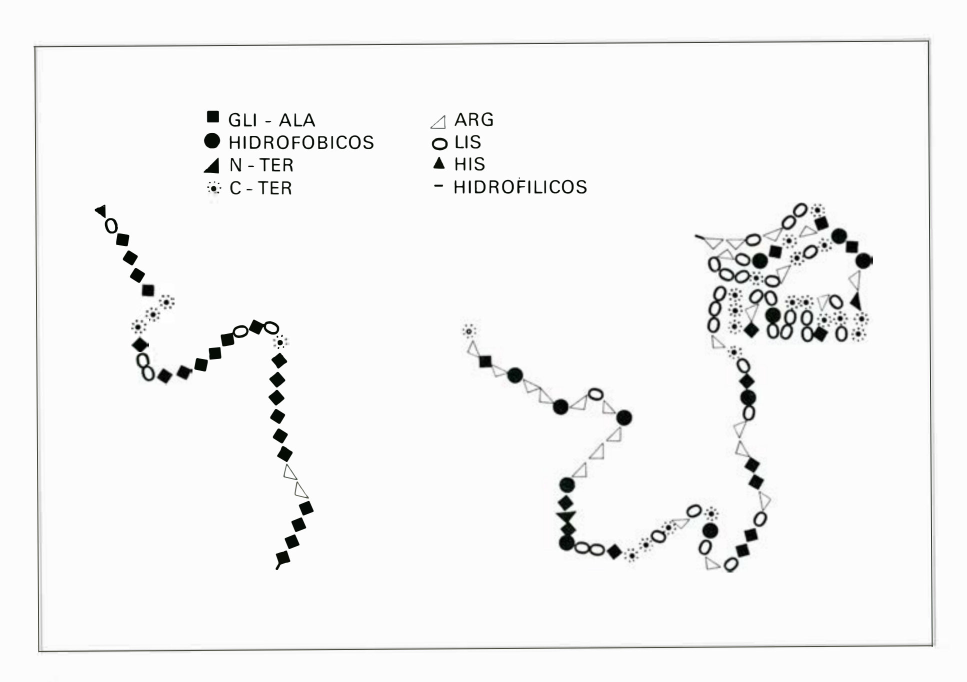
A partir de la aceptación de la existencia de grandes moléculas, vino una época de sistematización cuyo símbolo puede ser la labor del Profesor Paul Flory. En sus trabajos tomó como punto de partida la libertad de rotación de los enlaces estableciendo los conceptos fisicoquímicos básicos que relacionan el tamaño con la forma de los polímeros. Como resultado de esta libertad de rotación, las cadenas macromoleculares varían constantemente de forma en el espacio. Flory sistematizó las consecuencias que tiene este fenómeno básico sobre las propiedades de los polímeros, tanto fundamentales como aplicadas, llegando a establecer, por ejemplo, una teoría sobre la elasticidad del caucho de enorme interés práctico. Sin embargo, quedan todavía muchos problemas sin resolver en este campo. Uno de los más típicos es el que se denomina "volumen excluido". Para calcular las dimensiones de un polímero, con una cadena principal formada por átomos de carbono unidos uno a continuación del otro, hemos de tener en cuenta tanto la libertad de rotación de los enlaces como el hecho de que donde hay un átomo de carbono no puede estar otro, es decir, cada átomo implica un volumen del espacio que no puede ser ocupado por otro. Matemáticamente resulta difícil calcular cuándo, en su movimiento al azar, la molécula intentará pasar por un sitio que está ya ocupado. Este es un problema fisicoquímico que no ha podido resolverse de una forma definitiva. Podría pensarse que con los ordenadores actualmente disponibles sería un problema trivial calcular de cuántas maneras puede disponerse una macromolécula en el espacio y en qué casos intentará pasar dos veces por el mismo sitio. Para ver que esto no es así, imaginemos que a cada enlace carbono-carbono le damos únicamente tres posibilidades, las dos denominadas *gauche* y la *trans.* Así resultará que una cadena de 1000 átomos tiene $3^{1000} = 10^{477}$ posibles formas en el espacio. Podría pensarse que esta cifra es astronómica, pero en realidad es mucho mayor. Puede calcularse que el número máximo de partículas elementales (neutrones, electrones, protones) en todo el universo conocido es 10^{125}, es decir, que el número de posibles formas en el espacio de esta molécula resulta inimaginable y mucho más que astronómico. Si comparamos esta cifra con las posibilidades de un ordenador, sabiendo que necesita unos 20 nanosegundos aproximadamente para realizar una suma, tendremos que únicamente puede hacer 10^{16} sumas al año: un número ridículo comparado con las distintas formas que puede tener esta molécula en el espacio. Con ello queremos simplemente indicar que toda la físico-química de los polímeros está limitada



LA DIALECTICA ORDEN-DESORDEN de las macromoléculas, típica de todos los fenómenos de la naturaleza, se ilustra aquí con dos modelos moleculares. En el modelo superior, correspondiente al poli (acrilato de metilo), material semejante al plexiglás, se muestra un fragmento de dicha molécula. Esta conformación no es fija, sino que, debido a la libertad de giro de los enlaces, va cambiando continuamente su organización en el espacio. Según la temperatura, este cambio es más o menos pronunciado. A resultas de lo cual, el material se comporta en unas condiciones como un vidrio rígido, mientras que en otras se asemeja a un caucho. Muchos plásticos y todos los cauchos deben sus propiedades peculiares a esta posibilidad de cambio de forma. En el modelo inferior, correspondiente al nylon, se ilustra el caso opuesto. Las interacciones entre diferentes zonas moleculares producen una ordenación rígida, condicionada sobre todo por los puentes de hidrógeno *(blanco)* entre los átomos de oxígeno *(rojos)* y nitrógeno *(azul)*. Además, como característica adicional de ordenación, las moléculas se doblan sobre sí mismas de un modo regular, siguiendo un curso de zigzag tal como aparece en el modelo. Estas características son típicas de los polímeros cristalinos, que en general poseen una elevada resistencia mecánica, dando lugar a la mayoría de las fibras sintéticas y a muchos plásticos de interés técnico.



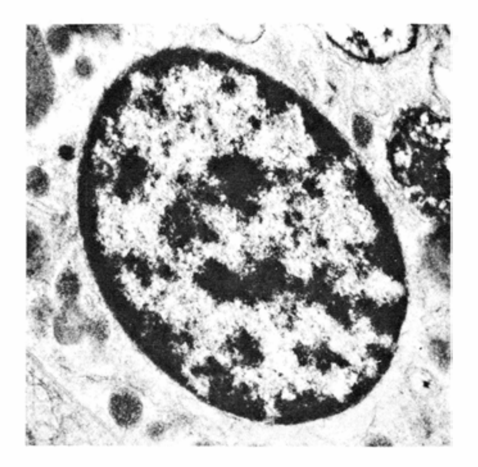
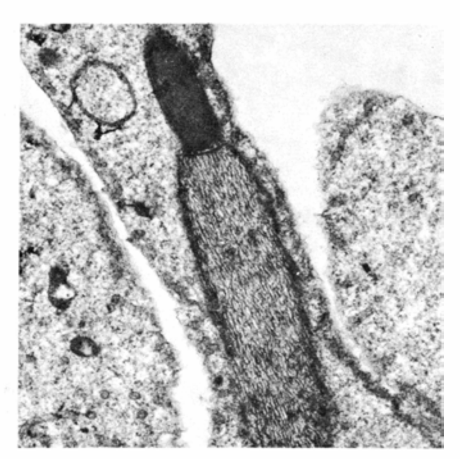
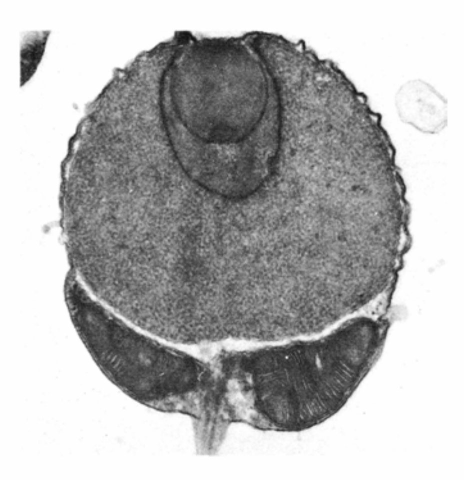
LOS POLIMEROS BIOLOGICOS poseen una ordenación espacial compleja, pero específica; es decir, todas las moléculas de un mismo tipo tienen la misma organización, debido a la interacción precisa entre los distintos grupos que la forman. Esta ordenación intramolecular contrasta con las características de los polímeros sintéticos que hemos mostrado en los modelos con que se encabeza este artículo. En la figura se muestra como ejemplo la estructura de la mioglobina, una proteína muscular. Debido a su arquitectura específica, esta proteína tiene la propiedad de almacenar una molécula de oxígeno, para suministrarla a la célula cuando así lo requiera su actividad metabólica.



LAS PROTEINAS BASICAS asociadas con el ADN en los distintos tipos de núcleos celulares pertenecen a dos familias principales: las protaminas y las histonas. Las protaminas se encuentran en el núcleo de los espermatozoides de un gran número de animales y se caracterizan por el elevado contenido en aminoácidos básicos. Sin embargo, su estructura química es muy variable en distintos organismos, variando tanto el tamaño como el tipo de aminoácidos presentes. Las histonas se encuentran en las células somáticas (hígado, cerebro, tejido muscular, etc.). Existen cinco clases fundamentales y su estructura química es prácticamente idéntica en todos los organismos, tanto animales como vegetales. Cuando están libres, la ordenación espacial de estas proteínas es pequeña, asemejándose a ovillos estadísticos. Sin embargo, al interactuar entre ellas y con el ADN, aparece una ordenación específica semejante a la de otras proteínas. En la figura se muestra un ejemplo de cada uno de estos dos tipos de proteínas: la salmina, una protamina procedente del salmón, y la histona H4.

por el hecho de que el cálculo estadístico de los promedios de propiedades y sus interacciones es una tarea poco menos que inasequible. De todas formas hay simplificaciones, y Flory ha tenido el mérito de derivar algunas propiedades por métodos estadísticos, pero éste es uno de los problemas que subsisten y que conceptualmente aparece como insoluble.

Junto a esta tendencia al cambio de forma, inducido por el movimiento molecular, las fuerzas interatómicas originan interacciones que pueden llegar a anular dicho movimiento molecular, apareciendo así estructuras muy ordenadas, como en el caso del nylon, ilustrado en la figura que encabeza este trabajo. En las macromoléculas naturales o biopolímeros, especialmente en los ácidos nucleicos y proteínas, esta ordenación es muy patente. La existencia de diferentes grupos químicos dispuestos ordenadamente en la molécula dan lugar a interacciones específicas. Como resultado, las proteínas normalmente no tienen una estructura en ovillo estadístico, sino que adoptan una forma precisa en el espacio, es decir, todas las moléculas de una misma proteína tienen la misma forma y sus grupos químicos ocupan siempre las mismas posiciones. En otras palabras, en los biopolímeros predominan las interacciones entre diferentes grupos, mientras que en los polímeros sintéticos predomina la libertad de rotación de los enlaces. Como resultado de esta forma espacial precisa, los biopolímeros, y especialmente algunas proteínas, tienen tres características peculiares: poder actuar como catalizadores específicos (enzimas), poder controlar esta catálisis y poder organizarse en estructuras más complejas. La primera característica, es decir, su capacidad de facilitar las reacciones químicas de la célula, se debe precisamente a su estructura espacial definida. En todas las proteínas con actividad enzimática, ésta reside en un punto de la proteína, denominado centro activo, en el que la orientación espacial de algunos grupos de la proteína es la adecuada para que allí se aloje una molécula determinada (substrato) cuya transformación química es favorecida en el entorno creado por el centro activo. La interacción entre los grupos proteicos y el substrato es de tal especificidad que así se consigue relajar los enlaces químicos del substrato, de modo que éste sea transformado químicamente con gran rapidez. Las proteínas son así capaces de actuar como catalizadores de una gran selectividad y eficacia. Una de las aspiraciones de los químicos es precisamente poder sintetizar

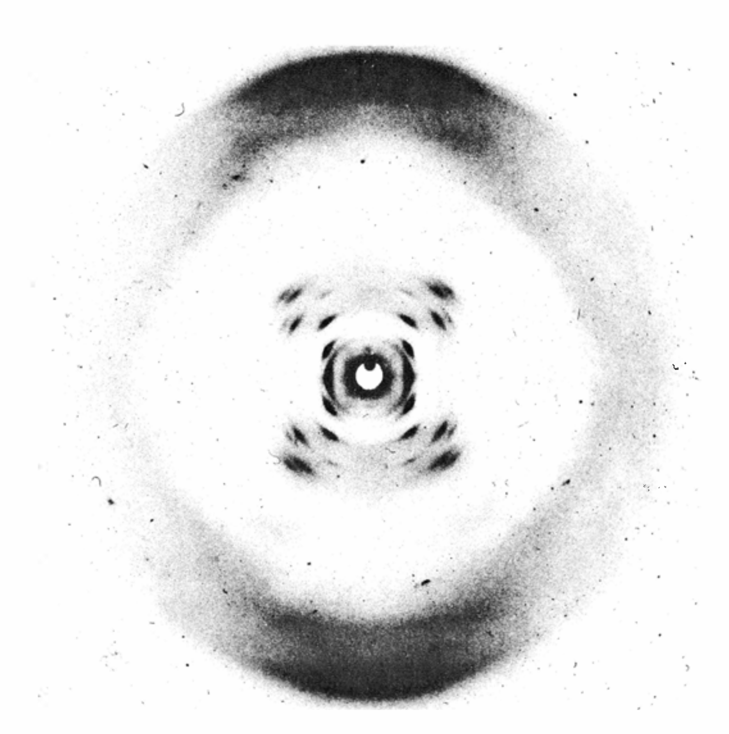
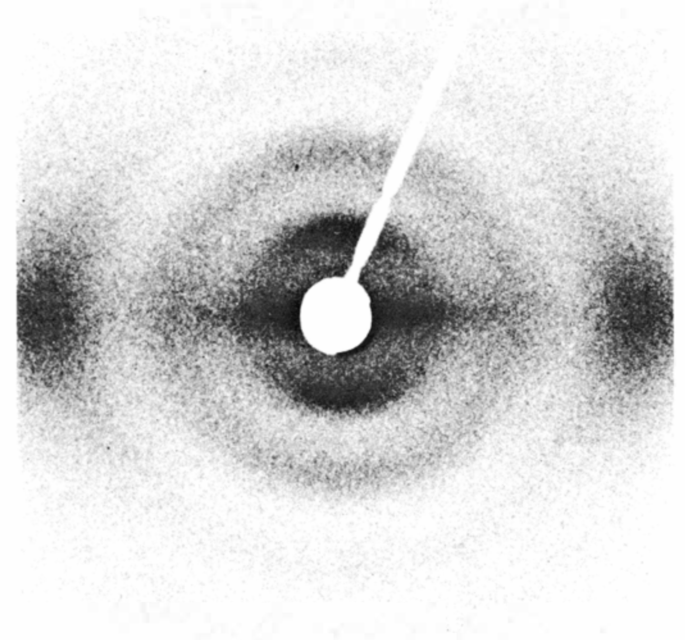
LOS COMPLEJOS ADN-PROTEINA que se encuentran en el núcleo celular tienen un aspecto variable en diferentes tejidos. En los núcleos somáticos, tal como el que se presenta a la izquierda, la cromatina aparece como una masa irregular, más o menos condensada según el tipo de tejido. En cambio, en los espermatozoides se observa una ordenación mucho más regular. En el centro se presenta una fotografía de un espermatozoide de pulpo en formación, pudiendo observarse que los complejos ADN-protamina se van ordenando en paralelo de una forma regular. A la derecha se presenta un espermatozoide de pepino de mar (*Holothuria tubulosa*), en el que la cromatina aparece como un ovillo de fibras de 250 angstroms de diámetro aproximadamente. El núcleo central contiene protaminas, mientras que los dos extremos contienen histonas. Las tres imágenes se han obtenido mediante un microscopio electrónico y los aumentos respectivos son: 13.000, 10.000 y 20.000 veces.

macromoléculas "a medida", que actúen como catalizadores específicos de procesos previamente determinados y es probable que este objetivo se alcance en fecha no lejana.

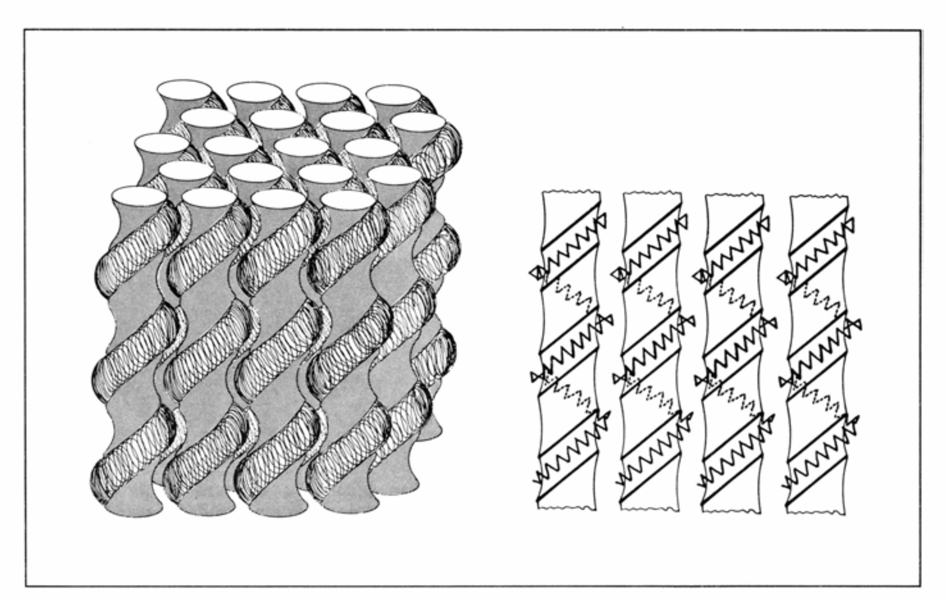
Otra característica peculiar de los catalizadores biológicos es que, debido precisamente a su estructura macromolecular, puede controlarse su actividad de acuerdo con las necesidades metabólicas de la célula. Una forma drástica de modificar su actividad es mediante sustancias inhibidoras, que pueden ocupar el centro activo e impedir que el enzima actúe. Otra forma, más sutil, de control puede producirse mediante el denominado efecto alostérico: algunas sustancias pueden interactuar con grupos determinados de la proteína, variando su ordenación espacial, de modo que el centro activo queda afectado en la orientación de sus grupos. Como resultado, la actividad catalítica del enzima es modificada, siendo éste uno de los mecanismos de control con que cuenta la célula para responder a los estímulos ambientales a que se encuentra sometida.

Acidos nucleicos y proteínas: el núcleo celular

Hay un tercer aspecto de los biopolímeros que es su posibilidad de organizarse en estructuras más complejas. La asociación de macromoléculas puede dar lugar a sistemas mucho más complicados,

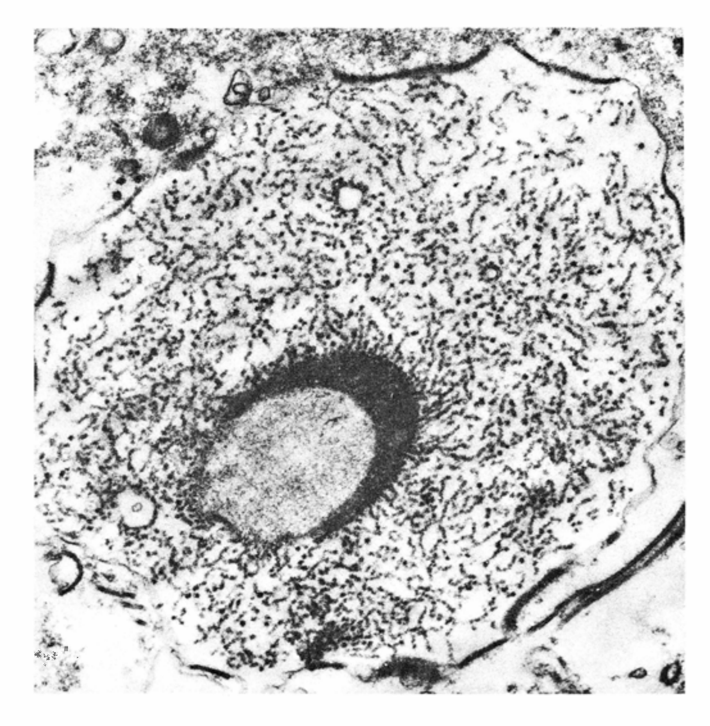
LA DIFRACCION DE RAYOS X permite determinar la organización espacial de los complejos formados por el ADN y las proteínas asociadas. En el diagrama de la izquierda, correspondiente a un complejo ADN-protamina, se observa una ordenación regular de la difracción, lo que indica un empaquetamiento también regular de las moléculas de ADN, tal como se representa esquemáticamente en la ilustración siguiente. En el diagrama de la derecha, correspondiente a un complejo ADN-histonas, se observa una serie de anillos, que indican cierto grado de orden, pero mucho menos preciso que en el caso de las protaminas. Los dos diagramas no están a la misma escala. El de nucleoprotamina corresponde a la zona de espaciados comprendidos entre 3 y 30 angstroms, mientras el de nucleohistona se sitúa entre 25 y 250 angstroms.

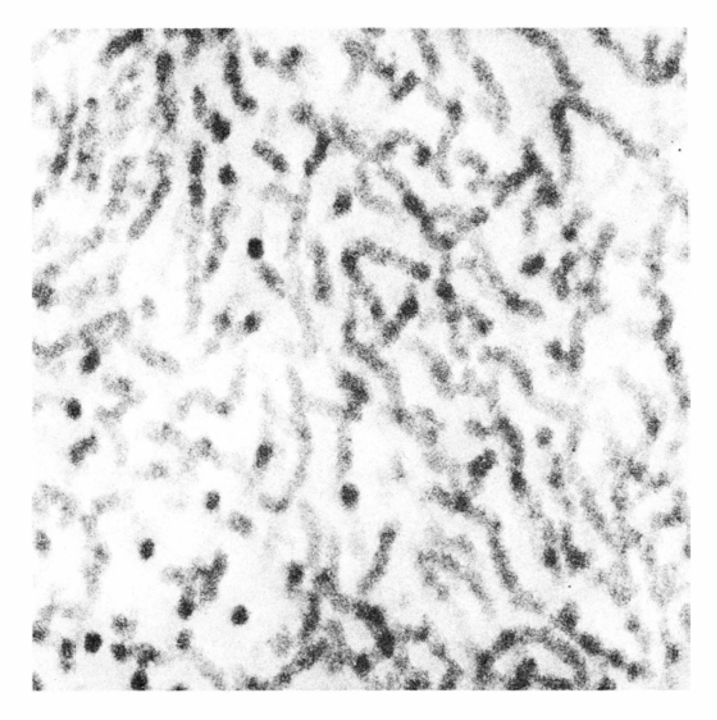


EN MUCHOS ESPERMATOZOIDES las moléculas de ADN se ordenan de una forma regular, casi cristalina. Las protaminas asociadas recubren el ADN siguiendo un surco a lo largo de su superficie helicoidal. El conjunto así formado se asemeja en cierta manera a un paquete de tornillos.

y puede decirse que cada uno de nosotros, desde el punto de vista químico, no es más que una asociación articulada de macromoléculas. Los complejos macromoleculares formados por los ácidos nucleicos y las proteínas son quizá los que tienen un mayor interés por estar a la base de los fenómenos vitales. Entre éstos podemos destacar los virus, los ribosomas y los complejos que forma el ADN con proteínas diversas en el núcleo celular y en los cromosomas. En las páginas que siguen vamos a discutir precisamente nuestro estado de conocimientos sobre este último tipo de complejos. No trataremos de los ribosomas, que han sido discutidos con frecuencia en las páginas de esta revista. Tampoco vamos a tratar de los virus, aunque hemos de remarcar que desde hace tiempo han atraído la atención de los investigadores por su simplicidad estructural, que permite estudiarlos con gran detalle. Normalmente están formados por una molécula de ácido nucleico protegida por una envoltura proteica constituida por unas pocas clases de proteínas distintas, y en algunos casos se trata de un solo tipo de proteína. A pesar de su simplicidad, el estudio de los virus ofrece un gran interés para comprender cómo dos tipos de macromoléculas distintos, el ácido nucleico y las proteínas, se asocian para dar lugar a una estructura más compleja, el virus. El estudio de las etapas de esta morfogénesis ofrece un interés fundamental para poder llegar a comprender los procesos morfogenéticos mucho más complejos que tienen lugar en los organismos superiores y, en definitiva, en el hombre.

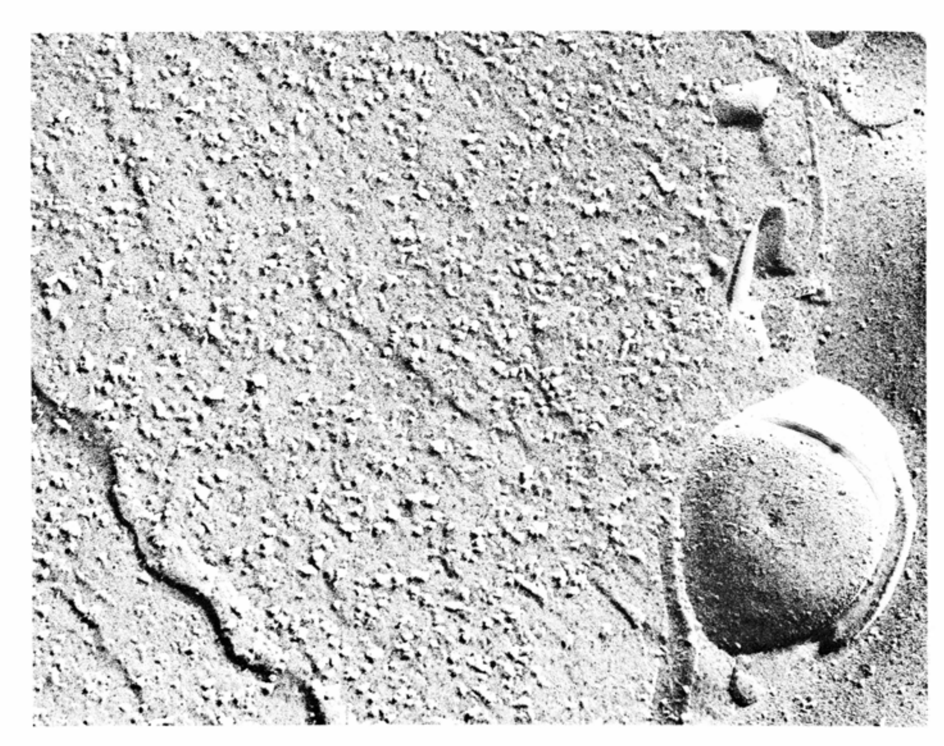
En contraste con la estructura relativamente simple de los virus, el núcleo celular ofrece mucho mayor complejidad. Desde un punto de vista estructural nos muestra aspectos muy variables, tal como puede apreciarse en la ilustración de la página 75. En las células somáticas, los complejos ADN-proteína, o cromatina, aparecen como masas irregulares más o menos uniformemente distribuidas por el interior del núcleo. Durante la mitosis estas masas se individualizan y adoptan la forma característica de los cromosomas. Por otra parte, en los espermatozoides, el núcleo adquiere una forma característica para cada especie y muy variable incluso en especies relativamente próximas, aunque en general puede decirse que la cromatina ofrece un aspecto mucho más compacto y uniforme que en las células





AL EXPLOTAR un núcleo de espermatozoide de holothuria, las fibras de cromatina que estaban empaquetadas se esparcen y puede estudiarse mejor su constitución. Así puede verse que estas fibras tienen un diámetro regular, alrededor de los 250 angstroms, y siguen trayectorias tortuosas. Sin embargo, se desconoce su estructura interna. La fotografía de la izquierda está aumentada 19.000 veces y la de la derecha 64.000 veces.

somáticas. El análisis químico muestra que la causa de estas diferencias externas reside en una diferencia notable en las proteínas asociadas con el ADN. En las células somáticas y en algunos tipos de espermatozoides, el ADN está acompañado por unas proteínas básicas denominadas histonas, de las que existen cinco tipos, casi idénticos en todos los organismos. Además hay otras proteínas denominadas acídicas, menos íntimamente asociadas al ADN y mucho más variables en su composición, de las que poco sabemos, pero que parecen desempeñar un papel muy importante en el control de la actividad genética, aunque su papel estructural parece ser poco significativo. En cambio, en los espermatozoides, las proteínas básicas presentes son muy variables en distintas especies, siendo las protaminas las más estudiadas de todas ellas.

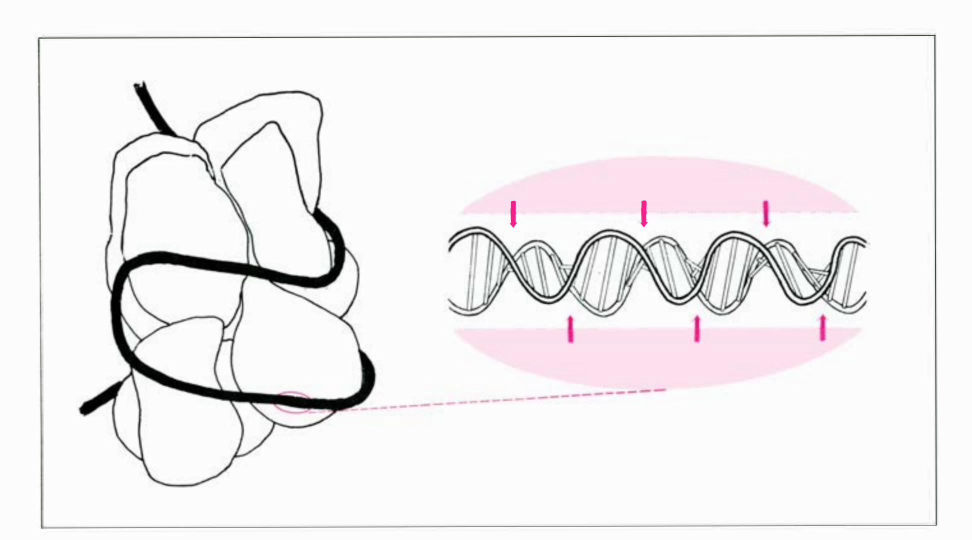


LOS NUCLEOSOMAS pueden visualizarse por algunas técnicas especiales de microscopía electrónica. En la fotografía se muestra parte de un núcleo de timo de ratón, aumentado 90.000 veces. Para obtener esta imagen, se rompió un fragmento de tejido congelado y se hizo una copia de la superficie de fractura así obtenida por evaporación de platino y carbono. En el núcleo se observan numerosas partículas, de 100 angstroms de diámetro aproximadamente, denominadas nucleosomas, y que son los constituyentes fundamentales de la cromatina. Sus características y propiedades han sido estudiadas en detalle por distintas técnicas bioquímicas y fisicoquímicas. A la derecha aparecen algunas vesículas de la membrana nuclear y del citoplasma, típicas en este tipo de células.

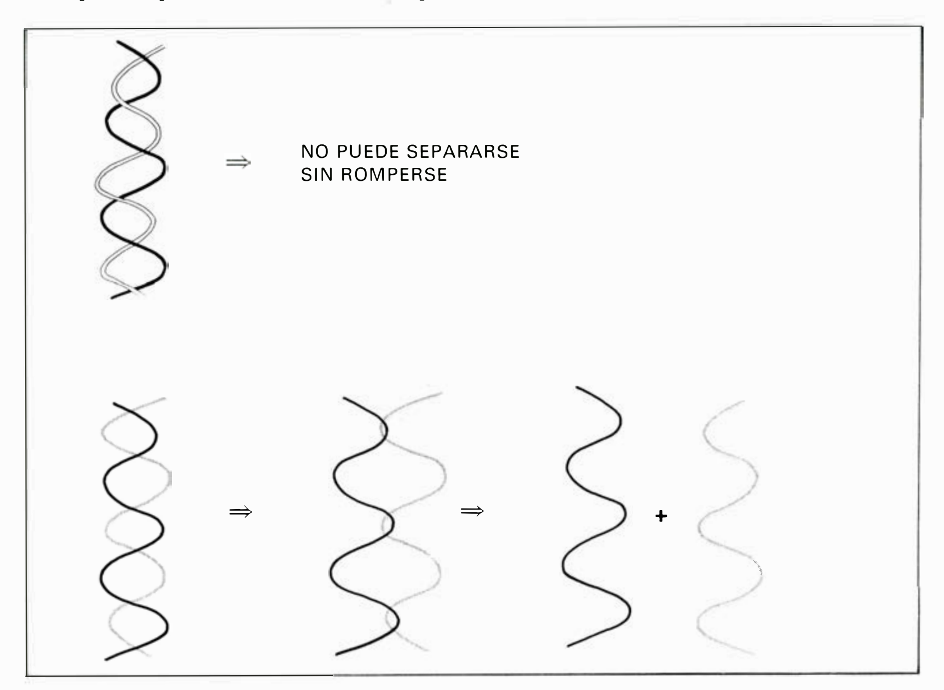
Las características estructurales de estas proteínas básicas están precisamente a la base de las diferentes ordenaciones que adopta el ADN y del aspecto variable que ofrecen los núcleos de los tejidos animales y vegetales. La salmina aislada, que es una protamina, se parece en su forma a un ovillo, aunque debido a sus cargas eléctricas tiene una estructura más rígida en forma de bastoncitos con poca flexibilidad, comparada con los polímeros sintéticos que comentábamos más arriba. Las histonas, en cambio, tienen una estructura más compleja. Según las condiciones en que se encuentren, zonas de la molécula adoptan una forma en ovillo, mientras otras zonas poseen una estructura bien definida; es decir, las histonas son un tipo intermedio de macromolécula, con una forma que oscila entre el ovillo y la estructura globular típica de la mayor parte de proteínas.

Las protaminas y las histonas interactúan con el ADN de una forma muy distinta, como puede apreciarse en los correspondientes diagramas de difracción de rayos X que tienen características muy distintas. En el caso de las protaminas, las moléculas de ADN se ordenan paralelamente de una forma muy regular, de manera que los complejos son prácticamente cristalinos y la conformación del ADN no sufre ninguna distorsión. Los núcleos de muchos espermatozoides son así pequeños cristales moleculares con una ordenación prácticamente perfecta. En cambio, los diagramas de difracción de los complejos ADN-histona muestran que el ADN no está bien ordenado y además aparecen anillos de difracción que indican una estructura repetitiva de un tamaño cercano a los 100 Å. Durante mucho tiempo se pensó que estos anillos de difracción indicaban que el ADN adoptaba una estructura en forma de superhélice, semejante a los arrollamientos múltiples de un filamento de bombilla. Sin embargo, recientemente, se ha podido establecer por métodos bioquímicos que la cromatina está formada por pequeñas partículas esferoides de unos 100 Å de diámetro denominadas nucleosomas. En efecto, se han encontrado enzimas capaces de romper la cromatina en partículas de este tamaño, que contienen ocho moléculas de histona de cuatro tipos distintos y un fragmento de ADN de unos 700 Å de longitud de alguna forma alojado en estas partículas. Los nucleosomas se han podido incluso visualizar utilizando técnicas adecuadas de microscopía electrónica. Este modelo permite explicar la falta de ordenación del ADN que aparece en los diagramas de difracción de rayos X, así como la existencia de una ordenación repetitiva de 100 Å de tamaño. Analizando estos diagramas puede llegarse a precisar que estas subunidades de los cromosomas están constituidas por un glóbulo de histonas rodeado por ADN que sigue una trayectoria complicada en el espacio. Esta trayectoria no está determinada con precisión, aunque parece que ha de tener una forma fija en cada nucleosoma.

Debe resaltarse que en este complejo de ADN e histonas hay un orden local, que si bien no sabemos exactamente cuál es, está perfectamente claro que se trata de un orden muy definido en la forma de asociación. Sin embargo, ignoramos totalmente la significación de esta arquitectura espacial tan definida. El ADN y las histonas están implicados en todos los procesos genéticos como constituyentes químicos del cromosoma, pero no sabemos cómo esta ordenación en nucleosomas puede mediatizar los procesos bioquímicos específicos en que está implicado el ADN: duplicación, síntesis de ácidos ribonucleicos, control de la actividad de un gene en diferentes células; y tampoco sabemos cómo estos nucleosomas se articulan entre sí para formar los cromosomas o el núcleo celular. Es decir, que aunque nosotros tenemos unos conocimientos estructurales más o menos bien definidos, no sabemos cómo relacionar la estructura con la función. Con carácter puramente especulativo en la ilustración de la página 78 se indica una posible función de estos nucleosomas que es totalmente heterodoxa dentro de los dogmas actualmente vigentes. El ADN, en su



EL NUCLEOSOMA está formado por ocho moléculas de histona alrededor de las cuales se dispone el ADN de una forma que no está aún determinada. Esta organización permite empaquetar el ADN en un volumen pequeño, pero es posible que su significación bioquímica sea mucho más amplia que un simple empaquetamiento. Una posibilidad sería que el ADN perdiera su conformación en hélice doble para adoptar otra conformación en la que las cadenas no estén arrolladas una en torno a la otra.



LA DUPLICACION DEL ADN se ve dificultada por su estructura en hélice doble. Las dos hebras no pueden separarse sin que se rompa la cadena o se produzcan rotaciones de la misma, que no parecen verosímiles dado el gran tamaño de las moléculas de ADN. De hecho, en las bacterias, en las que el ADN no está asociado a proteínas, se ha podido demostrar que la duplicación tiene lugar de un modo discontinuo, a base de pequeños fragmentos que van siendo después unidos entre sí. La organización en nucleosomas del ADN en los organismos superiores podría permitir una duplicación más simple. Si el ADN adoptara una conformación en la que las dos hebras no estuvieran arrolladas una en torno a la otra, podrían separarse fácilmente, tal como indica la figura. Quizás ésta es la significación más profunda de la asociación del ADN con proteínas para formar nucleosomas. Esta posibilidad pudo tener una significación evolutiva crucial en el paso de las bacterias a los organismos con núcleo capaces de almacenar una cantidad de ADN y una información genética mucho mayor.

replicación para dar dos moléculas hijas, tiene un problema topológico muy grave, ya que al estar las dos hebras en forma de hélice doble, una enrollada sobre la otra, toda la molécula ha de girar para poder separar las dos mitades, o romperse en pequeños fragmentos. Este es un problema que ha intrigado durante mucho tiempo a los bioquímicos. Si tenemos en cuenta que en el núcleo celular de las células superiores hay moléculas de ADN con un peso molecular del orden de miles de millones, el problema topológico de desenredar esta molécula puede ser muy grande. Puede así pensarse que la finalidad de estos nucleosomas fuese simplemente producir una desorganización local de la molécula de ADN, de manera que en lugar de estar en forma de doble hélice estuviera en forma de doble línea sinusoidal. Así, al tener dos hebras, una al lado de la otra, separarlas no presenta ningún problema topológico importante.

El descubrimiento de los nucleosomas ha supuesto un avance muy significativo en el conocimiento del núcleo celular. Sin embargo, queda aún un largo camino por recorrer. Desde un punto de vista estructural no sabemos cómo estos nucleosomas se asocian mutuamente. Si hacemos explotar un núcleo, la cromatina aparece como un conjunto de fibras sinuosas de un diámetro bastante regular, cercano a los 250 Å.

Ahora bien, ¿cómo se organizan los nucleosomas en el interior de estas fibras? Esta es una pregunta para la que no tenemos aún respuesta. Parece que hay un tipo de histona, la histona H1, que determina la forma de asociación de los nucleosomas, aunque su mecanismo de acción no está claro. Por otra parte, los diagramas de difracción indican que hay dos tipos de estructura en la cromatina, la globular, en forma de nucleosomas, y la fibrosa, en la que grupos de fragmentos de moléculas de ADN se sitúan paralelamente. La significación de este componente fibroso no está aún clarificada. A título de ejemplo en el diagrama de la página opuesta se indica una de las estructuras posibles.

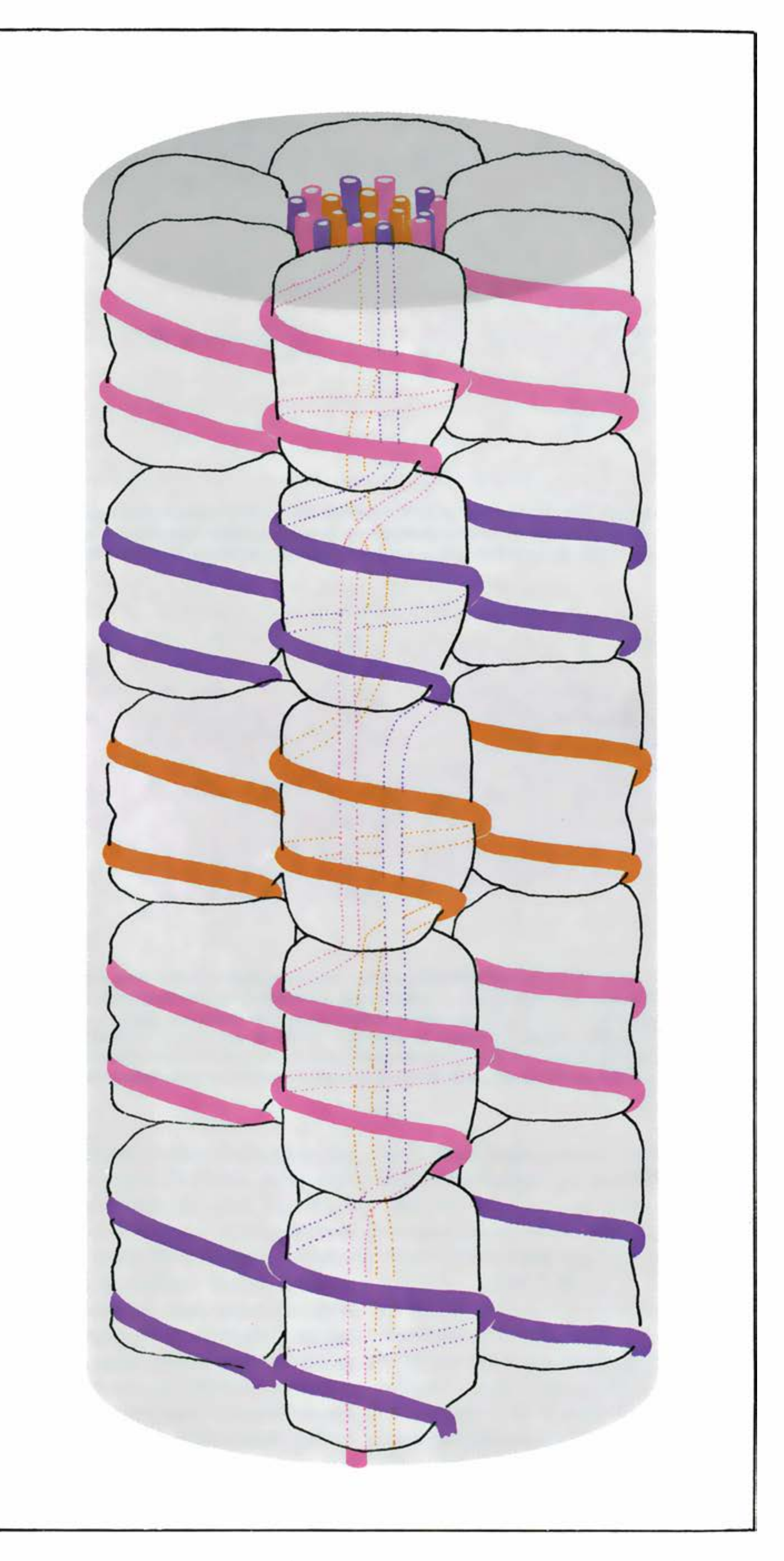
Podría pensarse que las técnicas de la microscopía electrónica serían útiles para poner de manifiesto la organización de los nucleosomas en el núcleo celular. Sin embargo, las posibilidades de este método son limitadas. En efecto, no existen aún técnicas aptas para observar núcleos hidratados, debido a que el microscopio electrónico trabaja en el vacío, condición en la que se evapora el agua, y la difracción de rayos X muestra que la presencia de agua es esencial para preservar la estructura de la cromatina. Por ello se ha intentado substituir el agua por materiales plásticos de forma que se mantenga la estructura de la cromatina, pero este objetivo aún no ha sido alcanzado. Además, en estas condiciones es difícil distinguir la cromatina del plástico en la que está embebida, y es preciso usar métodos de tinción que enmascaran la organización del ADN y de las proteínas. Por otra parte, al usar secciones muy finas, existe el peligro de que la elevada energía de los electrones usados en el microscopio destruya las moléculas que queremos observar. En definitiva, la microscopía electrónica se enfrenta con varios problemas técnicos y no ha sido aún posible determinar la estructura íntima de la cromatina a pesar de los esfuerzos desarrollados en esta dirección.

Como conclusión, y antes de pasar nuevamente a los polímeros sintéticos, podríamos decir que la situación actual en el estudio del material genético es la

siguiente: conocemos más o menos bien su estructura y tenemos técnicas para comprenderla, pero resulta muy difícil en el momento actual llegar a conocer su función. En especial se nos escapan los aspectos de control, es decir, cómo se regula la acción de estos complejos de macromoléculas, no sólo en este ejemplo que hemos descrito de los cromosomas, sino también en los ribosomas, membranas de la célula, sistema nervioso, etc. En este momento en que tenemos unos métodos analíticos bastante poderosos, es urgente llegar a una síntesis de nuestros conocimientos, es decir, en lugar de investigar aspectos parciales de las macromoléculas, llegar a estudiar los polímeros biológicos y sus complejos intactos. Parece que un retorno a la citología, a la observación de células individuales con los métodos analíticos de que ahora disponemos para estudiar con más detalle cada una de las proteínas, es un camino muy prometedor. Actualmente hay una gran efervescencia en este campo para estudiar, por ejemplo, la fusión entre células o la inyección de histonas en núcleos, es decir, aplicar unos métodos citológicos en células vivas que inicien la superación de las limitaciones que existen en el estudio de la biología a nivel molecular.

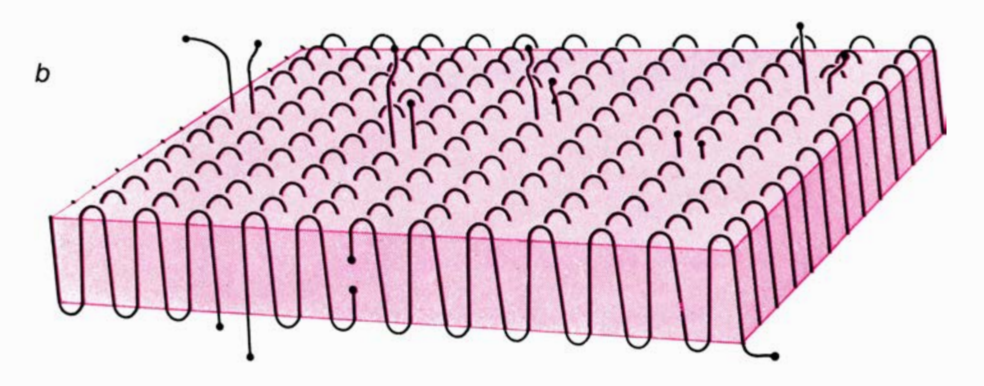
Perspectivas en el campo de los plásticos

A continuación revisaré brevemente cuál es, según mi criterio, la situación en el campo de los plásticos. Actualmente existe un progreso bastante rápido en cuatro campos en los que se han conseguido nuevos polímeros gracias al esfuerzo de gran número de químicos orgánicos, superando barreras que parecían prohibidas para conseguir nuevas síntesis. Estos campos de trabajo son: la resistencia al envejecimiento, la biodegradabilidad, la resistencia mecánica y la resistencia térmica. No vamos a tratar con detalle los dos primeros, sino que indicaremos únicamente que son dos procesos contrapuestos, en el sentido de que por un lado interesa encontrar polímeros que resistan a la intemperie, que no se deterioren, que mantengan sus propiedades mecánicas, y, por otro lado, interesa encontrar polímeros para uso corriente que sean fácilmente deteriorables en el ambiente exterior. De todas formas lo que se pretende en uno y otro campo no es lo mismo. Cuando se busca la resistencia al envejecimiento es en el caso de polímeros con una finalidad estructural, como puede ser la cubierta de una construcción, donde interesa disponer



LAS FIBRAS DE CROMATINA están formadas por nucleosomas, pero se ignora cuál es la disposición exacta de los mismos. En el esquema se indica una de las ordenaciones posibles, en el que las líneas de colores muestran la trayectoria seguida por el ADN. En la zona central aparecen moléculas de ADN paralelas, rodeadas de nucleosomas en una disposición cilíndrica. Para mayor claridad sólo se indica el ADN presente en algunos de los nucleosomas, quedando periódicamente fragmentos libres en el interior que forman un núcleo de moléculas paralelas. Esta organización es compatible con los diagramas de difracción de rayos X, que indican la existencia de dos tipos de ordenación en el ADN.

LOS PLASTICOS RESISTENTES A TEMPERATURAS ELEVADAS tienen estructuras químicas en las que predominan los anillos aromáticos. Algunos de estos materiales, como el indicado en último lugar, tienen además una gran resistencia mecánica debida a la rigidez de sus cadenas.



LOS CRISTALES DE POLIMEROS tienen una estructura característica cuyo elemento básico son lamelas delgadas, tal como la que se indica esquemáticamente en este diagrama. Estas lamelas tienen un espesor típico para cada material y temperatura. Las moléculas siguen un recorrido en zigzag en su interior, aunque los puntos de entrada y salida de una misma molécula en el cristal no han de estar necesariamente dispuestos de una forma tan regular como la indicada en la figura. Antes se mostró una molécula de nylon cuya organización espacial corresponde a esta estructura.

de un polímero con una gran resistencia al envejecimiento. La biodegradabilidad es una propiedad apreciada para los plásticos que se usan como envoltorios y envases, que son los que invaden nuestras playas. En este campo se están haciendo algunos progresos, se están explorando algunas ideas de combinar materiales degradables con otros que no lo sean, dando uno la resistencia y el otro una biodegradabilidad que aparece al cabo de un tiempo. De todas formas querría indicar que el problema ecológico de los plásticos quedaría muy simplificado si se eliminase uno de ellos, el polietileno. Efectivamente, este problema en las playas, mares y ríos es debido sobre todo a los plásticos que flotan en el agua, y entre los plásticos más corrientes (polietileno, policloruro de vinilo y poliestireno),el poliestireno y el policloruro de vinilo (PVC) se hunden y van degradándose lentamente hasta ser destruidos o acaso entran a formar parte de los sedimentos marinos. Pero, en cambio, el polietileno tiene la propiedad de flotar y, al ser un material de una resistencia química notable, es el que causa más molestias ecológicas y estéticas. No se trata de ningún problema fácil, pues el polietileno es precisamente el material que más se presta a su utilización en envases y envoltorios.

Otra limitación tradicional de los materiales plásticos es su escasa resistencia térmica. Al calentarlos se funden o descomponen a temperaturas moderadas, entre los 100 y 200° C normalmente. Por ello se ha trabajado activamente para sintetizar polímeros cuyas propiedades no se alteren al menos hasta alcanzar los 300° C. Para conseguirlo se ha intentado copiar la estructura del grafito, uniendo anillos uno tras otro, dando lugar a polímeros en forma de escalera, es decir, con dos cadenas paralelas constituidas por anillos aromáticos. Así se han conseguido algunas "escaleras" bastante largas, aunque en la práctica se ha visto que es posible conseguir una elevada resistencia térmica con polímeros, cuya estructura es mucho más simple. Estos ejemplos indican que la combinación de anillos aromáticos y la eliminación del enlace carbono-carbono en la cadena principal conducen a polímeros de una resistencia térmica muy superior a los plásticos normales. Al mismo tiempo se ha visto que algunos de estos polímeros son mucho más resistentes mecánicamente que los conocidos hasta ahora, consiguiéndose materiales con una resistencia mecánica, a igualdad de peso, notablemente superior al acero. Desgraciadamente su precio es muy elevado y sólo tienen aplicaciones restringidas para aquellos casos en los que es crítica una disminución en peso. Analizando las razones por las que aparece esta elevada resistencia mecánica, se ha visto que es debida a una elevada rigidez de las cadenas, lo que permite una cristalización de las moléculas en paralelo, sin el plegamiento típico que indicábamos más arriba con el ejemplo del nylon. En vista de este comportamiento nos podemos preguntar por qué en unos casos aparece un plegamiento de las cadenas y en otros no. ¿Qué factores condicionan la cristalización de los polímeros?

Para obtener la resistencia teórica máxima de un polímero, las moléculas han de situarse en paralelo. La resistencia a la rotura del enlace carbono-carbono o de cualquier enlace orgánico es muy elevada, de forma que si se consigue disponer las moléculas paralelas se alcanza una resistencia mecánica muy elevada y superior a la del acero. A primera vista parece que esta ordenación debería poder conseguirse en cualquier plástico, incluso en materiales de poco precio, tales como el polietileno. Sin embargo, los polímeros no cristalizan normalmente de esta forma, sino que tienden a cristalizar en láminas muy delgadas de un grosor de unos 100 Å, en las que las moléculas están plegadas en zig-zag. Esto hace que en la práctica no sea posible normalmente conseguir polímeros con una orientación paralela de todas las moléculas. Como consecuencia de esta estructura lamelar no puede conseguirse la resistencia mecánica teórica. La resistencia mecánica viene entonces condicionada por las fuerzas de cohesión existentes entre las lamelas, mucho más dé-

biles. En estas circunstancias, al sintetizarse polímeros resistentes térmicamente, en algunos casos se consiguieron moléculas muy rígidas, con la particularidad de no poder doblarse para formar lamelas, resultando polímeros que de un modo espontáneo se ordenaban en paralelo. Y aunque su heterogeneidad química hace que, en teoría, estas substancias sean menos resistentes que el polietileno, por ejemplo, en la práctica son las únicas que, de momento, se acercan a la resistencia mecánica teórica que puede tener un material plástico.

Frente a este comportamiento podemos preguntar cuál es la razón subyacente que obliga a los polímeros a cristalizar en forma de lamelas. La teoría más en boga interpreta este comportamiento como un fenómeno cinético, considerando que al cristalizar aparecen núcleos de una longitud determinada cuya velocidad de crecimiento es óptima para un espesor constante de las lamelas. No obstante, puede aventurarse también una explicación termodinámica. En efecto, en los cristales corrientes, de moléculas pequeñas, éstas tienen libertad para vibrar cada una en su lugar. En cambio, en un cristal de polímero, puede pensarse que las vibraciones coordinadas de las moléculas hagan preciso que exista una zona de relajación. Como resultado, para que el cristal sea estable se necesitará una zona de libertad de rotación a fin de que las vibraciones no destruyan el orden cristalino. La estructura en lamelas, limitando la longitud de las zonas de vibración coordinada y permitiendo una cierta libertad en los extremos, daría así mayor estabilidad a los cristales de polímeros. De todas formas esta explicación termodinámica no ha sido todavía elaborada en detalle, pero debe remarcarse que resolver si se trata de un fenómeno termodinámico o cinético tiene una gran importancia tecnológica: en efecto si se trata de un fenómeno termodinámico, es decir, si la estabilidad de los cristales determina un cierto grosor de las lamelas, no hay ningún otro camino posible para obtener polímeros totalmente orientados que sintetizar nuevos polímeros. Al contrario, si se trata de un fenómeno cinético, cabe pensar en la posibilidad de desarrollar métodos tecnológicos aptos para obtener polímeros orientados. El dogma actual es que nos encontramos ante un fenómeno de base cinética. Sin embargo, la interpretación termodinámica no puede excluirse. Aferrarse al dogma vigente podría ser un caso más en que el desarrollo tecnológico se vea frenado por conceptos erróneos.

"Segundos mensajeros" en el cerebro

Las células nerviosas se comunican entre sí por medio de la secreción de neurotransmisores. Estas sustancias constituyen mensajes químicos que, a través de unos segundos mensajeros, se traducen en varios efectos fisiológicos

James A. Nathanson y Paul Greengard

Para que un organismo pluricelular pueda sobrevivir y funcionar de forma eficaz, es necesario que sus células componentes actúen de un modo coordinado. Dicha coordinación requiere la transferencia de información entre células pertenecientes a partes del organismo que están separadas por distancias relativamente largas. En la mayoría de los animales superiores, existen esencialmente dos vías de comunicación intercelular: el sistema endocrino y el sistema nervioso. En el sistema endocrino, unas células especializadas secretan hormonas, las cuales son transportadas por el torrente sanguíneo hasta alcanzar zonas del cuerpo más alejadas, en donde influyen en la actividad de unas células diana de respuesta específica. En el sistema nervioso, una red de células nerviosas con finas prolongaciones establece comunicaciones intercelulares por medio de la secreción de sustancias neurotransmisoras. Estas atraviesan la breve hendidura que existe entre dos neuronas y modifican la actividad eléctrica de la célula receptora.

Ambos sistemas implican la acción de moléculas mensajeras que son liberadas por una célula, se desplazan cierta distancia y entran en contacto con la superficie de otra célula cuya actividad queda modificada. A la vista de esta semejanza, y suponiendo que la naturaleza funciona de la forma más económica, se puede postular que en algunos tipos de transmisión química entre células nerviosas pueden intervenir mecanismos similares a los que son mediadores de los efectos fisiológicos de determinadas hormonas. Este punto de vista ha conducido a nuevos enfoques de la organización bioquímica del cerebro y del mecanismo de acción de muchos fármacos que afectan al comportamiento.

A fin de proporcionar unas bases para discutir los mecanismos de la transmisión nerviosa, repasaremos primero algunos conceptos elementales referentes al mecanismo de acción hormonal. Las hormonas regulan una amplia gama de procesos bioquímicos en el interior de sus células diana, pudiendo actuar de tres formas: modificando la velocidad con que se producen enzimas y otras proteínas, afectando a la actividad de enzimas implicados en vías metabólicas clave o alterando la permeabilidad de las membranas celulares. Puesto que dichas acciones se llevan a cabo en el interior de la célula diana, debe haber allá alguna forma de percibir la propia hormona o la información que lleva.

Al parecer, la entrega de un mensaje hormonal al interior de la célula se puede realizar de dos maneras. Las hormonas esteroides, que son derivados del colesterol, se disuelven con facilidad en las grasas y, por tanto, pueden atravesar la membrana externa –esencialmente lipídica– de la célula diana, afectando directamente a los procesos que se dan en su interior. La cortisona y las hormonas sexuales estradiol y testosterona son ejemplos de este tipo de hormonas. En cambio, las hormonas derivadas de péptidos y de aminoácidos, tales como la insulina y la adrenalina (epinefrina), no pueden atravesar la membrana celular debido a su tamaño o a su estructura molecular. Lo que hacen es unirse a unos lugares receptores especializados, situados sobre la superficie de la célula, y ejercen su influencia sobre la maquinaria bioquímica desde el exterior.

Los receptores hormonales –grandes moléculas proteicas incrustadas en la membrana celular– son bastante selectivos en cuanto a su capacidad de unión con la hormona para la cual han sido diseñados, sin duda porque la configuración molecular del receptor permite que la molécula hormonal se encaje de forma bastante precisa. Las fuerzas que mantienen la unión entre la hormona y el receptor no son de tipo covalente, como las que unen entre sí los átomos de una molécula; son fuerzas más débiles que no tardan en liberar a la hormona, dejando al receptor suelto y preparado para recibir nuevas moléculas de la misma hormona. De lo anterior se deduce que el grado de alteración producido por una hormona en su célula diana depende de la concentración de la hormona en el líquido externo que baña a la célula y de la afinidad de la hormona por el receptor de la membrana.

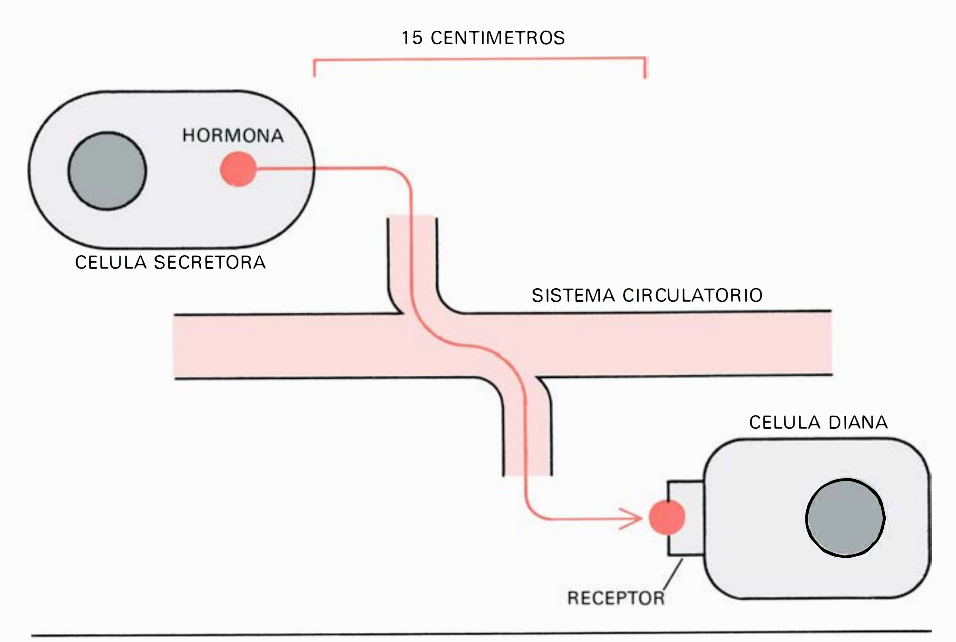
Una vez la hormona se ha unido al receptor, ¿cómo transmite su mensaje al interior de la célula? Los primeros en plantear esta cuestión, hace unos veinte años, fueron Earl W. Sutherland y sus colaboradores, de la Case Western Reserve University. Por aquel entonces, se hallaban estudiando el mecanismo por el que la hormona adrenalina provoca, por parte de las células del hígado, una liberación del azúcar glucosa al torrente sanguíneo cuando el organismo se encuentra en un estado de emergencia. La glucosa liberada procede de la hidrólisis del glucógeno (almidón animal), que se conserva almacenado en el hígado. Sutherland y su colega Theodore W. Rall comprobaron que, cuando se ponían en presencia de adrenalina membranas celulares aisladas a partir de células hepáticas, se producía un factor no identificado, el cual, al ser mezclado con el citoplasma de las células hepáticas, imitaba la acción de la adrenalina, provocando la conversión de glucógeno en glucosa. Por tanto, la respuesta a la hor-

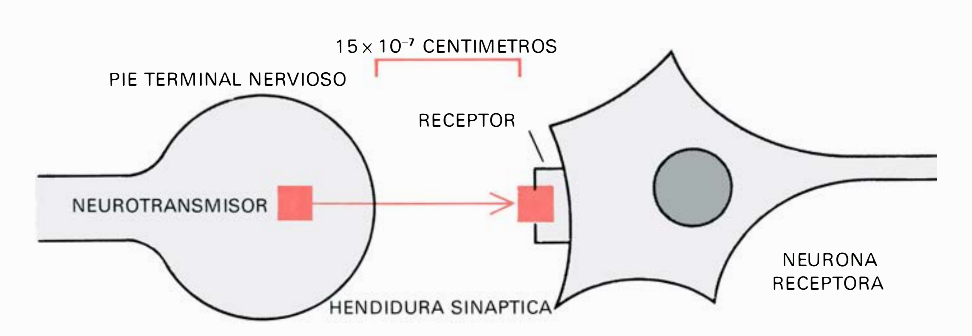
mona parecía presentar dos etapas: la interacción de la hormona con la membrana resultando la formación de un factor no identificado, seguida por la activación, por parte del factor, del mecanismo bioquímico del citoplasma.

Experimentos posteriores identificaron al factor como monofosfato de adenosina cíclico (AMP cíclico), que pertenece a un grupo de pequeñas moléculas denominadas nucleótidos. Su estructura está relacionada con la del trifosfato de adenosina (ATP), que es la forma universal en que se encuentra la energía química de la célula. El atributo "cíclico" del AMP cíclico se refiere al hecho de que el único grupo fosfato (PO_4) de la molécula forma un anillo con los átomos de carbono a los que está unido.

Poco después, Sutherland y sus colaboradores demostraron que las membranas de las células hepáticas (y de otras muchas células) contenían un enzima, la adenilciclasa, que transformaba el ATP en AMP cíclico. Dado que el ATP se encuentra casi exclusivamente en el citoplasma, Sutherland y sus colegas G. Alan Robinson y Reginald W. Butcher llegaron a la conclusión teórica de que al menos una parte de la molécula de adenilciclasa debía estar orientada hacia el interior y que el AMP cíclico producido por la acción del enzima debía liberarse en el interior de la célula. Al exponer células hepáticas en presencia de adrenalina, observaron que se da un aumento sustancial de la velocidad de conversión de ATP en AMP cíclico por acción de la adenilciclasa, poniendo así de manifiesto la existencia de una relación funcional entre la unión receptor-hormona en la cara externa de la membrana y la activación de la adenilciclasa en la cara interna. Posteriormente, se demostró que dicha vinculación se da en una amplia variedad de hormonas que se unen a receptores de la membrana; ahora bien, habida cuenta de que los distintos tipos de células poseen receptores para hormonas diferentes, una hormona dada incrementará los niveles de AMP cíclico en sus células diana, pero no en otras células.

Hoy en día se acepta de un modo general que el AMP cíclico, generado por la adenilciclasa como respuesta a la unión de una hormona al receptor de la membrana, actúa como un "segundo mensajero" que desde la membrana retransmite el mensaje de la hormona (que es el primer mensajero) a la maquinaria bioquímica de la célula. De esta forma, la señal de bajo nivel, que representa la hormona, puede ser amplificada miles de veces mediante la fabricación de AMP cíclico.





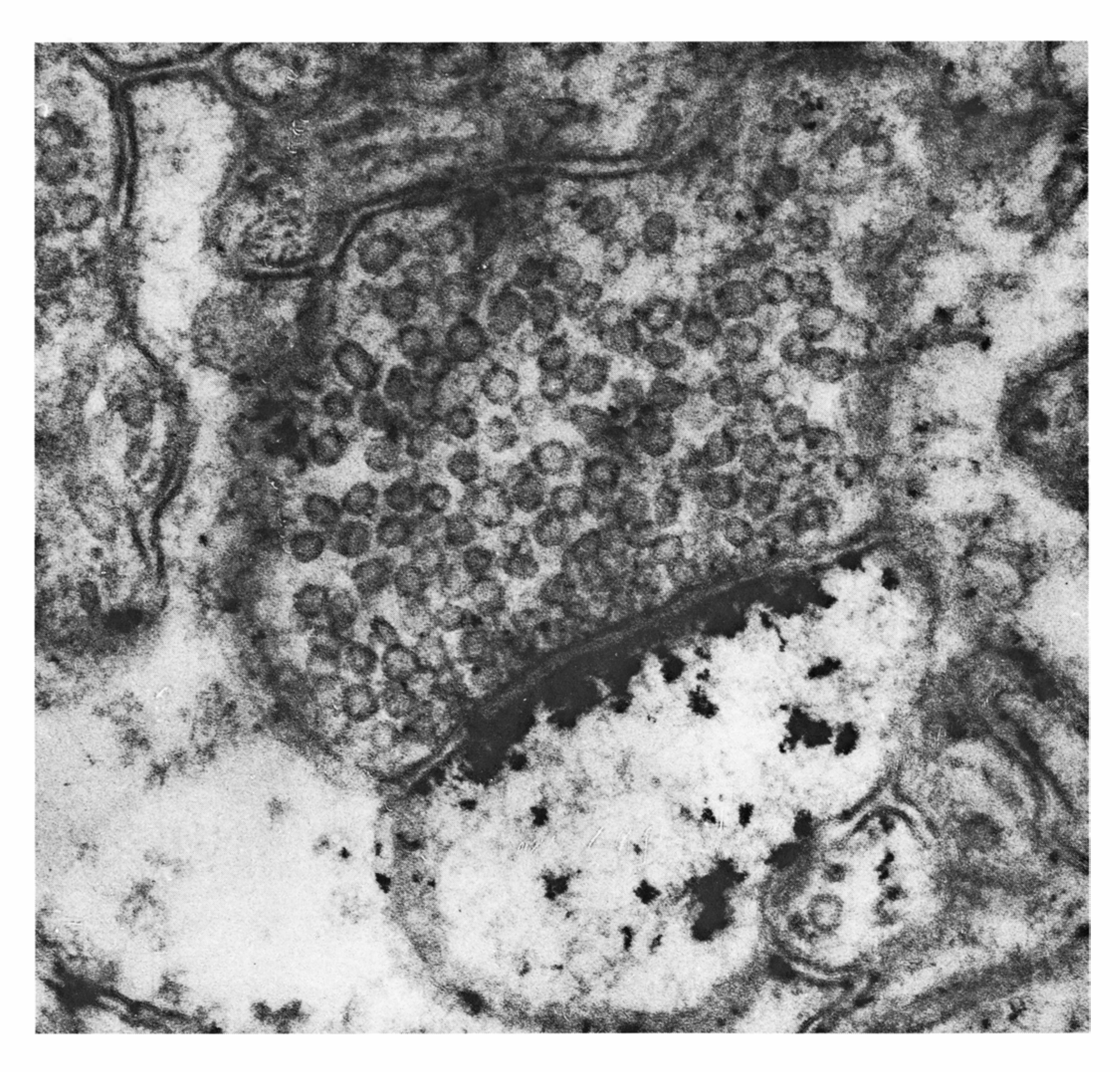
MECANISMOS ANALOGOS, que están en la base de la comunicación en el sistema endocrino *(arriba)* **y en el sistema nervioso** *(abajo)*. **En ambos mecanismos hay una célula que libera mensajeros químicos (que son hormonas y neurotransmisores respectivamente), los cuales atraviesan el espacio extracelular y se unen a receptores de la superficie de otra célula, modificando así la actividad de esta última. Al parecer, los efectos intracelulares de un gran número de hormonas y de neurotransmisores se dan por mediación de un segundo mensajero: el AMP cíclico.**

Hasta ahora, hemos considerado las interacciones que se dan entre hormonas, receptores y AMP cíclico en células que no pertenecen al sistema nervioso. ¿No podrían unos mecanismos análogos ayudar a explicar el modo en que se realiza la comunicación entre células nerviosas? La célula nerviosa, o neurona, es la unidad estructural y funcional básica del sistema nervioso central; el resultado neto de la interacción compleja de muchas neuronas constituye el comportamiento. La información se transmite a lo largo de la fibra alargada de una neurona bajo la forma de un impulso electroquímico. No obstante, el impulso se detiene cuando alcanza la minúscula sinapsis, o hendidura, que separa el pie terminal de la fibra de su correspondiente neurona receptora. A fin de que pueda ser saltada la sinapsis y se pueda volver a iniciar la transmisión electroquímica, el neurotransmisor es liberado y atraviesa la hendidura que media entre una célula y la siguiente. Al igual que las hormonas que hemos estado considerando, los neurotransmisores no penetran en la célula receptora, sino que entran en contacto con los receptores de la cara externa de la membrana celular, produciendo un cambio del potencial eléctrico entre el exterior y el interior de la membrana.

La primera vez que se relacionó el AMP cíclico con el cerebro fue cuando Sutherland y sus colaboradores hallaron grandes cantidades del enzima adenilciclasa en el cerebro de animales vertebrados, lo cual indicaba que el cerebro era un lugar de síntesis activa de AMP cíclico. En 1967, Eduardo De Robertis, en colaboración con Butcher y Sutherland, observó que, cuando se disgrega el tejido cerebral por homogeneización y los componentes subcelulares resultantes se separan según su densidad mediante un proceso de centrifugación, las

fracciones que contienen una mayor cantidad de partículas constituidas por pies terminales son también las que presentan niveles de actividad más altos para la adenilciclasa. Este hecho constituía un hallazgo interesante, ya que muchas de las partículas mencionadas comprenden fragmentos de membrana procedentes de ambos lados del contacto sináptico y representan, por tanto, las áreas concretas del cerebro en las que se realiza la comunicación entre células nerviosas. El hecho de que la adenilciclasa estuviera asociada a dichas áreas de un modo específico sugería que el AMP cíclico podría desempeñar algún papel en la transmisión sináptica.

En estos mismos experimentos, se comprobó que las partículas constituidas por pies terminales también presentaban altos niveles de fosfodiesterasa, el enzima que se encarga de degradar al AMP cíclico para dar una forma de monofosfato de adenosina fisiológicamente inactiva. Poco tiempo después, Noel T. Florendo y Russell J. Barrnett, en colaboración con nuestro laboratorio de la escuela de medicina de la Universidad de Yale, desarrollaron un método citoquímico para determinar, con la ayuda del microscopio electrónico, la localización de la fosfodiesterasa en células aisladas. El uso de esta técnica les per-

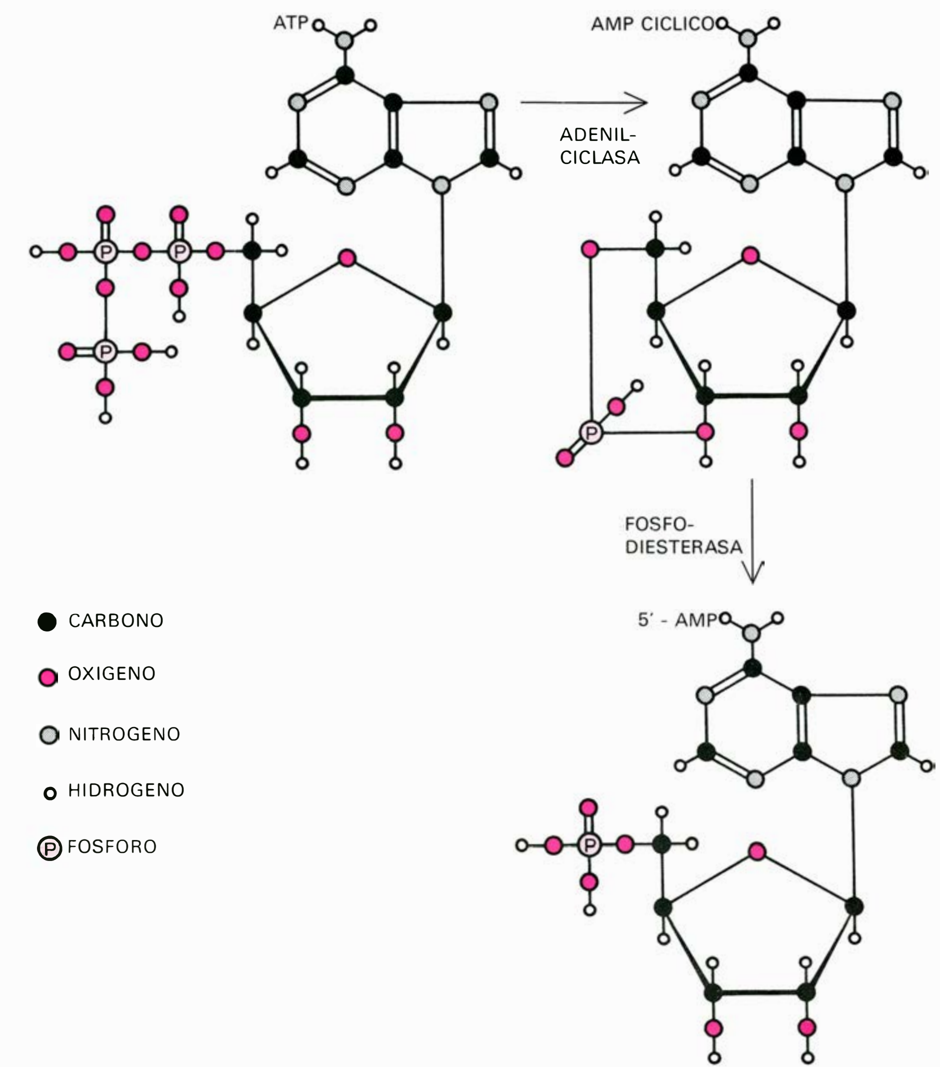


EL PAPEL DE SEGUNDO MENSAJERO que desempeña el AMP cíclico en la comunicación entre neuronas fue sugerido, en parte, porque se halló fosfodiesterasa, enzima encargado de degradar al AMP cíclico, en el contacto sináptico entre dos neuronas, es decir, en el lugar donde se da la transmisión nerviosa. Noel T. Florendo y Russell J. Barrnett, en colaboración con los autores, identificaron el enzima aplicando al tejido cerebral varios compuestos químicos que reaccionaban con el producto de la actividad enzimática formando un precipitado denso. Luego observaron el tejido al microscopio electrónico. En la microelectrografía que aquí se presenta, se observa un solo contacto sináptico en el que el precipitado indicador de la situación de la fosfodiesterasa está asociado a una región de la membrana de la célula receptora que posee receptores de sustancias neurotransmisoras. Este hallazgo aporta pruebas a favor de la hipótesis de que las moléculas de neurotransmisor que son expulsadas de las vesículas circulares del pie terminal nervioso *(arriba)* se unen a los receptores e inducen la fabricación de AMP cíclico, siendo éste el mediador de diversas acciones intracelulares posteriores. La sinapsis registrada se halla aumentada unas 138.000 veces.

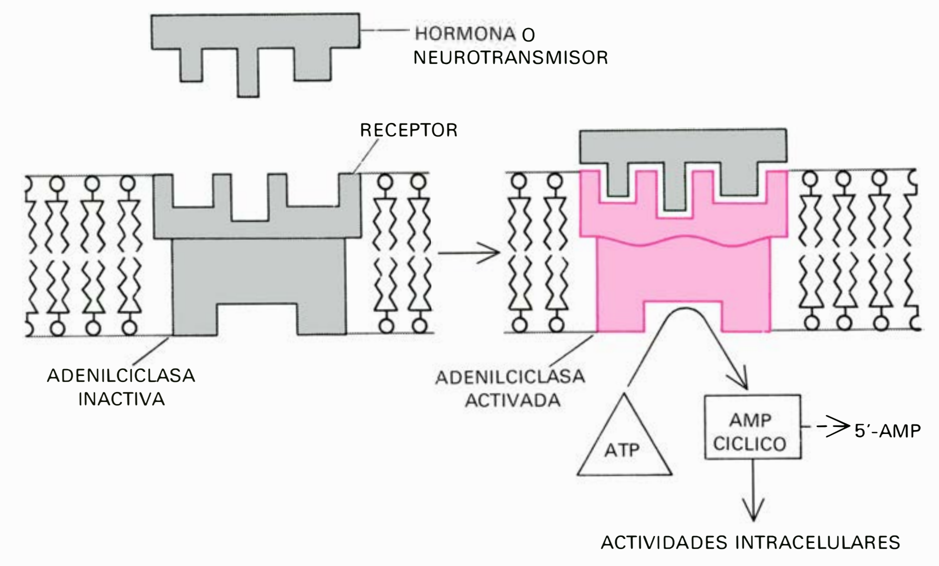
mitió demostrar que la actividad fosfodiesterásica de la región sináptica se halla localizada en una zona de la membrana de la célula receptora que, según se cree, sería la portadora de receptores de sustancias neurotransmisoras. Dicha situación del enzima era un hecho más que abogaba por un papel del AMP cíclico en la transmisión sináptica, ya que la fosfodiesterasa postsináptica tendría acceso, y por tanto podría degradar, a las moléculas de AMP cíclico previamente sintetizadas como resultado de la estimulación de la adenilciclasa por el neurotransmisor.

Más o menos al mismo tiempo que se realizaron los experimentos anteriores, otros trabajos demostraron que se podían conseguir considerables incrementos de los niveles de AMP cíclico en el tejido nervioso estimulándolo eléctricamente o con un neurotransmisor. Donald A. McAfee y Michel Schorderet estuvieron trabajando en nuestro laboratorio con un ganglio del sistema nervioso simpático localizado en el cuello, y, utilizando conejos y vacas, demostraron que la estimulación eléctrica de los nervios que inervaban al ganglio, cuyo resultado era la transmisión sináptica, estaba vinculada a una elevación de los niveles de AMP cíclico. Posteriormente, McAfee y John W. Kebabian observaron que, al aplicar el neurotransmisor dopamina a estos ganglios, el efecto producido –elevación de los niveles de AMP cíclico– imitaba al de la estimulación eléctrica y que, a la inversa, la aplicación de AMP cíclico era capaz de reproducir algunos de los efectos electrofisiológicos de la dopamina. Durante el mismo período de tiempo, Shiro Kakiuchi y Rall, de la Case Western Reserve University, y algo más tarde John W. Daly y sus colaboradores, del National Institute of Arthritis, Metabolism and Digestive Diseases, informaron de que los fragmentos resultantes de trocear tejido cerebral presentaban grandes incrementos de su contenido de AMP cíclico cuando se exponían en presencia de soluciones que contenían diversos neurotransmisores, como la norepinefrina o la histamina.

Aunque todos los experimentos anteriores indicaban la existencia de algún tipo de asociación entre los receptores del neurotransmisor y los niveles de AMP cíclico, dejaban sin resolver la siguiente cuestión: ¿podían los efectos de un neurotransmisor determinado estar acoplados directamente a la activación de la adenilciclasa? A fin de poder dar pruebas de la existencia de un vínculo funcional entre la unión de un neuro-



LA SINTESIS Y LA DEGRADACION del AMP cíclico son tareas que llevan a cabo dos enzimas que se hallan asociados a la membrana externa de determinadas células receptoras de hormonas o de neurotransmisores. La adenilciclasa convierte la molécula portadora de energía adenosintrifosfato (ATP) en AMP cíclico, por separación de dos grupos fosfato de su molécula y por unión posterior del grupo fosfato restante al esqueleto hidrocarbonado de la molécula, formándose así un anillo. La fosfodiesterasa inactiva al AMP cíclico: ello lo consigue al abrir el anillo formado por el grupo fosfato y transformar la molécula del nucleótido en una forma de AMP inerte.



LA ACTIVACION DE LA ADENILCICLASA se realiza cuando una hormona o un neurotransmisor se une a su receptor específico situado sobre la membrana celular. De inmediato, el enzima utiliza una parte del ATP que se encuentra presente en el citoplasma y lo pasa a AMP cíclico, el cual sirve para retransmitir la señal en cuestión desde la membrana hasta el interior de la célula.

transmisor determinado con su receptor y la síntesis ulterior de AMP cíclico, se tenía que demostrar la presencia de una adenilciclasa sensible a un neurotransmisor específico, es decir, un enzima cuya actividad dependiera en gran parte de la presencia de un neurotransmisor determinado. Antes de pasar a describir con cierto detalle este enzima, vale la pena que comentemos brevemente la relación que existe entre el receptor de un neurotransmisor y la acción de determinados fármacos que afectan al comportamiento.

Puesto que las células nerviosas se comunican básicamente mediante la liberación de neurotransmisores en los contactos sinápticos, es lógico deducir que cualquier cosa que interfiera la unión de un neurotransmisor con su receptor provocará una desorganización en la comunicación normal entre neuronas, alterando así el comportamiento del organismo. La causa de este tipo de interferencias puede ser la presencia de sustancias extrañas, como, por ejemplo, los fármacos, introducidas en el torrente sanguíneo. Si un fármaco dado tiene una configuración molecular semejante a la de un neurotransmisor endógeno, es posible que sea capaz de unirse al receptor de la membrana específico del neurotransmisor y de producir la misma acción que la sustancia endógena. Los fármacos de este tipo se denominan activadores de los receptores. Por otra parte, si el fármaco tiene una configuración similar a la del neurotransmisor pero la semejanza no es tan acentuada como en el caso de los activadores, es posible que sea capaz de unirse al receptor pero sin llegar a activarlo. En este caso, el fármaco impide que el receptor pueda ser activado por un neurotransmisor. Los fármacos de este tipo se llaman inhibidores de los receptores. Hay otros fármacos que afectan al comportamiento y no son ni activadores ni inhibidores. Algunos pueden presentar una acción combinada, uniéndose al receptor del neurotransmisor y produciendo una activación parcial; otros pueden afectar al receptor de una forma indirecta, alterando la cantidad de neurotransmisor disponible, es decir, capaz de unirse. En el trascurso de estos últimos años, este concepto relativamente sencillo, que describe las relaciones entre fármacos, receptores y comportamiento, ha constituido la base del considerable progreso en el conocimiento y y el tratamiento de los trastornos del sistema nervioso, y ha centrado la atención en el posible papel del AMP cíclico tanto en la función normal del cerebro como cuando ésta es deficiente.

Por ejemplo, se suele aceptar que los síntomas de la enfermedad de Parkinson son el resultado de la degeneración de un grupo de neuronas que se encuentran en una región denominada substantia nigra, situada en la base del cerebro, y cuyas fibras se proyectan sobre los ganglios basales situados cerca del centro del cerebro. Estas neuronas secretan el neurotransmisor dopamina en los pies terminales de sus fibras; su degeneración reduce la cantidad de dopamina disponible y, por tanto, capaz de entrar en contacto con los receptores de las

LA ADENILCICLASA SENSIBLE A LA DOPAMINA es el enzima que actúa como mediador entre el neurotransmisor dopamina y sus efectos postsinápticos; también es el lugar de acción de algunos fármacos que afectan al comportamiento. En esta ilustración se observa que la dopamina liberada por la actividad de los pies terminales presinápticos atraviesa la hendidura sináptica y se une a un receptor de dopamina de la membrana postsináptica, unión que provoca la activación de la adenilciclasa y, consecuentemente, la síntesis de AMP cíclico en la célula postsináptica. La actividad de la adenilciclasa sensible a la dopamina puede ser influenciada por fármacos que, según se sabe, se unen de forma específica a los receptores de dopamina. Los fármacos inhibidores, tales como el agente antiesquizofrénico clorpromacina (Toracina), bloquean al receptor e impiden su activación por la dopamina. Los activadores, la apomorfina por ejemplo, actúan de forma contraria, o sea: activan al receptor y provocan efectos similares a los de la dopamina.

células postsinápticas situadas en los ganglios basales. En consecuencia, las células postsinápticas receptoras empiezan a no funcionar normalmente y provocan los síntomas característicos de la enfermedad de Parkinson: temblores, rigidez y una demora en la iniciación del movimiento. Aunque en el cerebro de los pacientes que sufren dicha enfermedad la dopamina esté agotada, al parecer, no están dañados los receptores de dopamina situados en los ganglios basales.

Estos hechos, en gran parte puestos en evidencia por Arvid Carlsson, de la Universidad de Göteborg, Oleh Hornykiewicz, de la Universidad de Viena, y el desaparecido George C. Cotzias, del Brookhaven National Laboratory, han conducido a un nuevo tratamiento revolucionario para la enfermedad de Parkinson: la administración del fármaco *levo*-dihidroxifenilalanina (L-DOPA), que es el aminoácido precursor de la dopamina. Cuando se administra L-DOPA oralmente, penetra en el torrente circulatorio y es transportada hasta el cerebro, en donde se absorbe y se transforma en dopamina. (Ni siquiera el propio neurotransmisor puede penetrar en el cerebro desde la sangre.) De esta manera, la dopamina recién elaborada puede actuar como un activador: puede estimular los receptores de dopamina de los ganglios basales. La L-DOPA, al ser capaz de compensar la falta de dopamina endógena, puede eliminar algunos de los síntomas de la enfermedad.

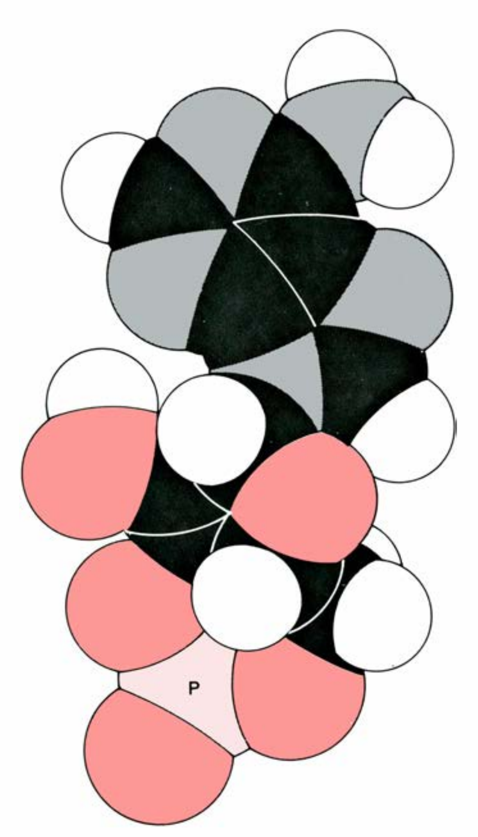
Los fármacos que actúan como inhibidores del receptor de dopamina también tienen una utilidad terapéutica. Un ejemplo de este tipo de fármacos es el tranquilizante fenotiacínico clorpromacina (Toracina), ampliamente usado en el tratamiento de la esquizofrenia. A menudo, los pacientes psicóticos tratados con clorpromacina sufren una mejoría notable en sus síntomas mentales, pero los efectos secundarios del fármaco limitan su utilidad en clínica. Así, tras un tratamiento prolongado, puede ocurrir que los pacientes empiecen a manifestar temblores y otros movimientos anormales, semejantes a los observados en la enfermedad de Parkinson. Cuando cesa la administración del tranquilizante, suelen desaparecer los movimientos anormales. Por tanto, parece ser que la clorpromacina lleva consigo la inducción farmacológica de la enfermedad de Parkinson por medio de un bloqueo de los receptores de dopamina situados en los ganglios basales, imitando así los síntomas debidos al agotamiento de la dopamina, aun cuando el neurotransmisor se halle presente en cantidades normales.

La enfermedad de Parkinson farmacológicamente inducida, aparte de ser un efecto secundario problemático en los tratamientos farmacológicos antipsicóticos, ha servido para arrojar luz sobre las anormalidades bioquímicas que podrían estar en la base de la esquizofrenia. Si los fármacos que parecen bloquear los receptores de dopamina reducen los síntomas de la esquizofrenia, quizá resulte que la esquizofrenia es la consecuencia de una actividad excesiva de las neuronas sintetizadoras de dopamina en algunas regiones del cerebro. El exceso de actividad de estas células produciría un exceso, también, de liberación de dopamina a partir de sus pies terminales, lo que llevaría a una estimulación excesiva de los receptores postsinápticos de dopamina. La clorpromacina, al bloquear estos receptores, impediría, según parece, esta sobrecarga de estímulos disminuyendo así los síntomas de la esquizofrenia.

El descubrimiento de estas dos aplicaciones terapéuticas (de los activadores del receptor de dopamina para luchar contra la enfermedad de Parkinson y de los inhibidores del mismo receptor para luchar contra la esquizofrenia) puso en evidencia la necesidad apremiante de una investigación destinada a un conocimiento más profundo de la naturaleza bioquímica del receptor de dopamina en el cerebro de los mamíferos. Los trabajos de Kebabian, McAfee y Schorderet, anteriormente expuestos, habían sugerido que el AMP cíclico desempeñaba posiblemente un papel importante como mediador de las respuestas de las neuronas a la dopamina. Sin embargo, para demostrar que existía una relación funcional entre la unión del neurotransmisor y la síntesis de AMP cíclico, había que identificar una adenilciclasa que sintetizara AMP cíclico a partir de ATP en presencia de dopamina.

En 1972, Kebabian, junto con Gary L. Petzold, de nuestro laboratorio, obtuvo pruebas experimentales demostrativas de la presencia de una adenilciclasa sensible a la dopamina en el núcleo caudado, uno de los ganglios basales del cerebro y una región rica en receptores de dopamina. El enzima estaba localizado en las membranas sinápticas y presentaba una similitud marcada con el receptor de dopamina. Su actividad podía ser estimulada con concentraciones de dopamina muy bajas e inhibida fuertemente por dos tipos de fármacos antiesquizofrénicos conocidos por su efecto de bloqueo del receptor de dopamina. Todos estos resultados daban bases firmes para suponer que el receptor de dopamina situado en el núcleo caudado y en algunas otras áreas del cerebro de los mamíferos sería, en realidad, un componente de la adenilciclasa sensible a la dopamina y que, por tanto, el AMP cíclico sería el mediador de la acción intracelular de la dopamina en algunas sinapsis.

Esta hipótesis ha sido ampliamente confirmada por unas investigaciones más profundas realizadas por Yvonne Clement, Cormier y Kebabian en nuestro laboratorio y por Leslie L. Iversen y sus colaboradores en la Universidad de Cambridge. Examinaron un gran número de sustancias y, en conjunto, se observó un paralelo significativo entre la capacidad para actuar como activador o inhibidor del receptor de dopamina y la capacidad de activar o inhibir la



MODELO ESPACIAL del AMP cíclico, que indica que se trata de una molécula compacta. Dicho nucleótido consta de un azúcar de cinco átomos de carbono, al que están unidos un anillo de adenina *(arriba)* y un grupo fosfato.

adenilciclasa sensible a la dopamina. Desde entonces, esta correlación ha llevado al desarrollo de una nueva metodología para una determinación rápida, y estudio posterior, de la potencialidad de los fármacos como activadores (agentes contra la enfermedad de Parkinson) o inhibidores (agentes contra la esquizofrenia) del receptor de dopamina. Mientras que los métodos tradicionales de medida de la actividad del receptor de dopamina requerían pruebas de comportamiento interminables y a menudo de poca precisión, la determinación de la actividad de la adenilciclasa sensible a la dopamina en presencia del fármaco que es objeto de la investigación proporciona una vía rápida y cuantitativa para evaluar la capacidad del fármaco, bien como bloqueador, bien como activador del receptor de dopamina.

Otra relación entre el AMP cíclico y la transmisión sináptica se puso de manifiesto con los trabajos de Floyd E. Bloom, Barry Hoffer y George Siggins, en el National Institute of Mental Health. Estos investigadores hallaron pruebas fisiológicas de que algunos efectos del neurotransmisor norepinefrina implicaban la mediación del AMP cíclico. Lo que entonces estaban estudiando Bloom y sus colegas era la regulación, por parte de la norepinefrina, de la actividad nerviosa del cerebelo, zona que controla muchos de los movimientos automáticos del cuerpo, como por ejemplo la marcha. Los elementos clave que regulan dichos movimientos son unas neuronas grandes y sumamente ramificadas, denominadas células de Purkinje, que están situadas en la corteza, o capa más externa, del cerebelo. Al extender el marco de los estudios anatómicos hechos por Tomas Hokfelt y Kjell Fuxe en el Instituto Carolino de Estocolmo, Bloom y sus colegas comprobaron que el locus coeruleus, que es un pequeño grupo de neuronas sintetizadoras de norepinefrina localizado en la profundidad del tallo cerebral, emite unas fibras delicadas que lo ponen en conexión con la corteza del cerebelo, y que la estimulación de esta vía de conducción provoca una marcada disminución de la frecuencia de emisión de impulsos por parte de las células de Purkinje. Al administrar, a las células de Purkinje aisladas, cantidades conocidas de norepinefrina o de AMP cíclico mediante una micropipeta, observaron que la aplicación de cualquiera de los dos compuestos producía una reducción de la frecuencia de emisión de impulsos nerviosos por parte de dichas células, actuando, por tanto, del mismo modo que la estimulación del locus coeruleus. Por otra parte, la utilización de una técnica de marcaje con sustancias fluorescentes, que teñía al AMP cíclico de forma selectiva, les permitió demostrar que tanto la estimulación del locus coeruleus como la aplicación directa de la norepinefrina provocaban una disminución muy considerable de los niveles de AMP cíclico en el interior de las células de Purkinje. A partir de los resultados anteriores y de otros experimentos complementarios, se ha emitido la hipótesis de que los efectos de la norepinefrina sobre la actividad nerviosa de las células de Purkinje se llevan a cabo a través de la estimulación de una adenilciclasa sensible a la norepinefrina y la consiguiente síntesis de AMP cíclico en el interior de la célula.

Unos experimentos recientes realizados en nuestro laboratorio parecen indicar que el papel del AMP cíclico en el funcionamiento de los receptores de las sustancias neurotransmisoras no está en absoluto restringido al sistema nervioso de los vertebrados. Por ejemplo, hemos hallado un enzima en los ganglios torácicos de los insectos que es capaz de actuar como mediador de los efectos del neurotransmisor serotonina. Esta adenilciclasa sensible a la serotonina puede ser activada mediante concentraciones muy bajas de serotonina; los fármacos conocidos como bloqueadores de los receptores de serotonina inhiben de forma específica el efecto de la serotonina sobre la actividad del enzima. Constituye un hecho interesante el que uno de los más potentes de estos agentes bloqueadores sea precisamente el alucinógeno dietilamida del ácido lisérgico (LSD). Estos hallazgos hacen pensar que el receptor de la serotonina localizado en el tejido nervioso está íntimamente asociado a una adenilciclasa sensible a la serotonina y que algunos de los efectos fisiológicos del LSD podrían ser el resultado de la inhibición del enzima.

Hoy en día se han encontrado un mínimo de cinco neurotransmisores capaces de estimular adenilciclasas específicas para los neurotransmisores respectivos: dopamina, norepinefrina, serotonina, histamina y octopamina. Por otra parte, se han obtenido pruebas recientes que sugieren la participación de otro nucleótido cíclico, el monofosfato de guanosina cíclico (GMP cíclico), como mediador de los efectos del neurotransmisor acetilcolina en algunas sinapsis, a través de una activación de la guanilciclasa, que es el enzima encargado de pasar el trifosfato de guanosina (GTP) a GMP cíclico. También cabe la posibilidad de que el GMP cíclico desempeñe algún papel como mediador de los efectos de la norepinefrina y de la histamina en determinados receptores, distintos de los asociados al sistema del AMP cíclico.

Cabe destacar que muchos de los fármacos que afectan al comportamiento por su acción sobre los receptores de las sustancias neurotransmisoras producen efectos semejantes a los que se presentan de forma natural en pacientes con enfermedades mentales o neurológicas. Por ejemplo, tanto los mineros expuestos a un envenenamiento por manganeso como los pacientes tratados con fármacos antiesquizofrénicos presentan un síndrome imposible de distinguir del de la enfermedad de Parkinson apareci-

NEUROTRANSMISOR	TIPO DE RECEPTOR
DOPAMINA	dopamina
NOREPINEFRINA	alfa-adrenérgico
	beta-adrenérgico
SEROTONINA	serotonina
HISTAMINA	H_1
	H_2
ACETILCOLINA	muscarínico (lento)
	nicotínico (rápido)
ENCEFALINA	opiáceo

LOS FARMACOS que inciden en el comportamiento se clasifican en dos: los que actúan por contacto directo con el receptor del neurotransmisor y los que afectan al receptor indirectamente, alterando la cantidad de neuro-

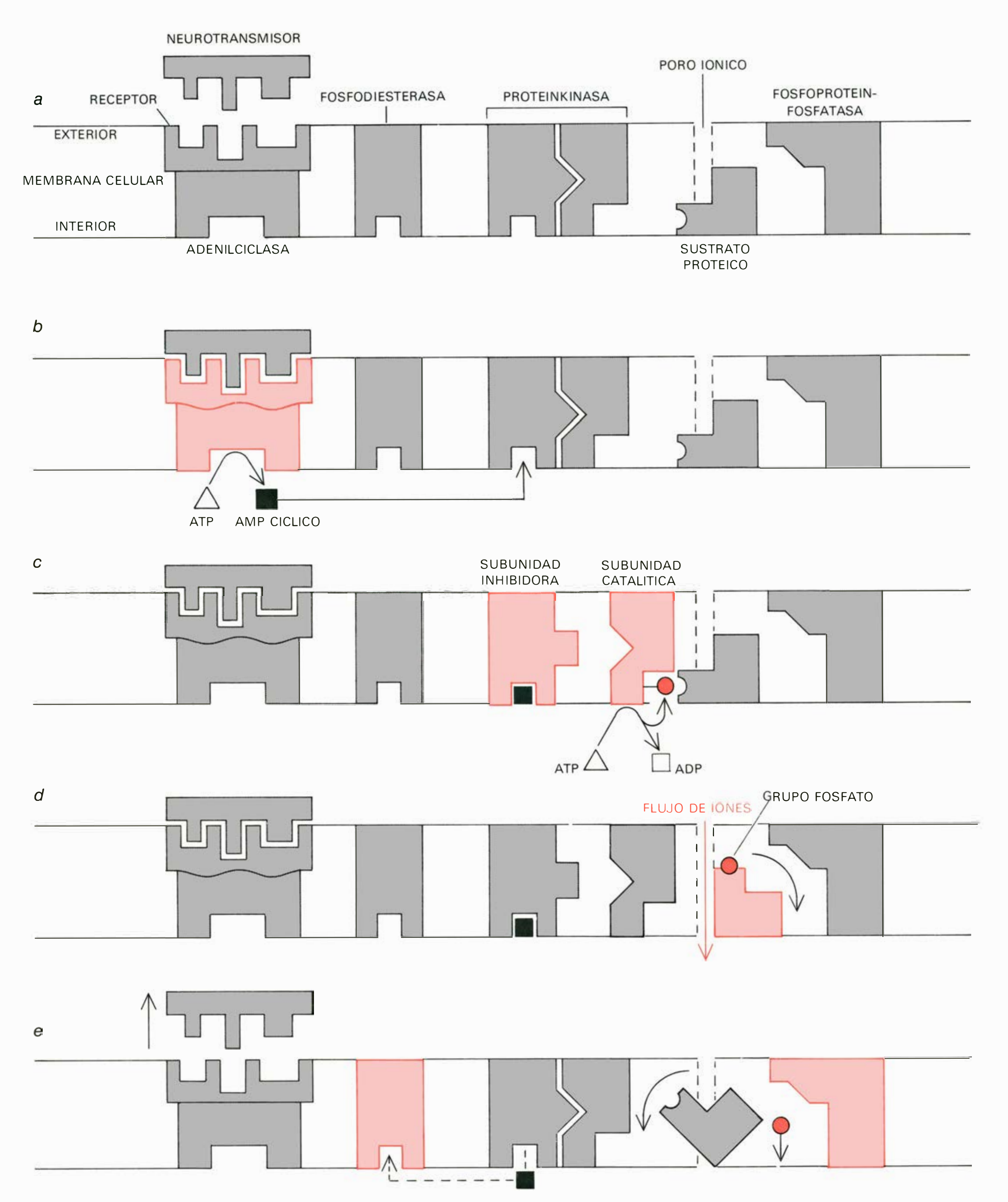
da de forma natural; por otra parte, el LSD puede producir alucinaciones similares a las experimentadas por los esquizofrénicos. El conjunto de estas observaciones clínicas, entre otras muchas, nos hace plantear la posibilidad de que algunas enfermedades neurológicas y mentales sean el resultado de anormalidades inherentes a los sistemas específicos receptor-adenilciclasa. Este tipo de anormalidades podrían tener un origen genético debido a mutaciones no letales o adquirirse por exposición a elevados niveles ambientales de determinadas toxinas, como puede ser el manganeso.

No obstante, no se ha observado que todos los receptores de sustancias neurotransmisoras estén asociados a la síntesis de un nucleótido cíclico. Por ejemplo, uno de los mecanismos sinápticos que mejor se conoce –el desencadenamiento de la contracción del músculo voluntario mediante la liberación de acetilcolina en el contacto entre la terminación nerviosa y el músculo– parece operar independientemente de la síntesis de un nucleótido cíclico (véase "La respuesta a la acetilcolina", de Henry A. Lester; INVESTIGACIÓN Y CIENCIA, abril). Ello no carece de sentido si se tiene en cuenta el hecho de que la participación de nucleótidos cíclicos en la transmisión sináptica implica una compleja serie de etapas –etapas que describiremos más adelante– cuya duración es relativamente larga dentro de la escala de tiempo característica de los acontecimientos a nivel de neuronas. No es de extrañar, por tanto, que aquellos tipos de transmisión en los que la velocidad es lo que priva, tales como la contracción del músculo voluntario, estén basados en mecanismos receptores más rápidos.

Hasta ahora hemos descrito el mecanismo por el que algunos neurotransmisores, al unirse a un receptor específico, estimulan la producción de AMP cíclico en el interior de la célula receptora. Pero aún queda por resolver una cuestión: ¿Cómo puede un incremento de los niveles de AMP cíclico transformar el mensaje del neurotransmisor en una acción fisiológica?

PRUEBAS DEL PAPEL MEDIADOR DE UN NUCLEOTIDO CICLICO		FARMACOS QUE ACTUAN A NIVEL DEL RECEPTOR DEL NEUROTRANSMISOR		FARMACOS QUE AFECTAN LOS NIVELES DE NEUROTRANSMISOR DISPONIBLE PARA UNIRSE AL RECEPTOR	
AMP CICLICO	GMP CICLICO	ACTIVADORES (DEL RECEPTOR)	INHIBIDORES (DEL RECEPTOR)	AUMENTAN LOS NIVELES	DISMINUYEN LOS NIVELES
sí	no	apomorfina alfa-bromcriptina	antiesquizofrénicos (Toracina, Haldol)	*levo*-dihidroxifenilalanina (L-DOPA) amantadina anfetaminas (Dexedrina) metilfenidato (Ritalin)	alfa-metil-para-tirosina
no	sí (?)	fenilefrina (Neosinefrina)	fentolamina (Regitina)	antidepresivos tricíclicos (Elavil, Tofranil) inhibidores de la MAO (Parnate) anfetaminas cocaína	reserpina (Serpasil) alfa-metil-dopa (Aldomet)
sí	no	epinefrina (adrenalina) isoproterenol (Isuprel)	propanolol (Inderal)		
sí	no	5-metoxi-N,N-dimetiltriptamina	dietilamina del ácido lisérgico (LSD) metisergida (Sansert)	antidepresivos tricíclicos (Elavil, Anafranil) triptófano	para-cloro-fenilalanina
no	sí (?)	2-metil-histamina	difenilhidramina (Benadryl) dimenhidrinato (Dramamina)	histidina amodiaquina	alfa-hidracinohistidina brocresina
sí	no	betazola (Histalog)	metiamida cimetidina		
no	sí (?)	pilocarpina carbacol betanecol (Urecolina)	escopolamina atropina (belladonna) propantelina (Probantine)	insecticidas (Paratión, Malatión) diisopropilfosforofluorhidrato (DFP) piridostigmina (Mestinon)	toxina de *Botulinus*
no	no	nicotina	d-tubocurarina (curare) succinil-colina (Anectina)		
sí	?	morfina heroína metadona meperidina (Demerol)	naloxona (Narcan)		

transmisor libre y, por tanto, disponible para la unión. Los fármacos que alteran los niveles de neurotransmisor actúan de formas variadas; algunos aumentan la síntesis de un neurotransmisor determinado (por ejemplo, la L-DOPA), otros provocan la liberación del neurotransmisor previamente elaborado (anfetamina), unos terceros bloquean la hidrólisis del neurotransmisor o su captación de la hendidura sináptica (ciertos antidepresivos). EL AMP cíclico y un compuesto relacionado, el GMP cíclico, parecen ser los mediadores de los efectos de varios neurotransmisores, por lo que también desempeñan un papel en la acción de un gran número de fármacos que afectan al comportamiento.



ESTE MODELO TEORICO explica cómo un incremento de los niveles de AMP cíclico en el interior de la neurona receptora puede dar lugar a un cambio pasajero de la permeabilidad iónica de la membrana celular, alterando así la excitabilidad eléctrica de la célula. En la sección *a* se presentan de forma muy esquemática los diversos componentes moleculares cuyo papel en el proceso ha sido determinado: convencionalmente, la membrana se halla en su estado de reposo. Cuando un impulso nervioso alcanza el pie terminal presináptico, se libera el neurotransmisor. Este atraviesa la hendidura sináptica y se une a los receptores de la célula receptora, produciendo una activación de la adenilciclasa que convierte ATP en AMP cíclico *(b)*. El AMP cíclico así formado se une a la subunidad inhibidora de la proteínkinasa, provocando la disociación de ésta y dejando libre la subunidad catalítica, es decir, la forma activada del enzima, que transfiere un grupo fosfato del ATP a un sustrato proteico *(c)*. La adición del grupo fosfato produce un cambio de forma o de posición del sustrato proteico, cambio que permite el flujo de determinados iones a través de los poros que perforan la membrana y altera la excitabilidad eléctrica de la célula *(d)*. El cese de la reacción comienza cuando el neurotransmisor se disocia del receptor, terminando, por tanto, la síntesis de AMP cíclico. la fosfodiesterasa inactiva las reservas restantes. Finalmente, la fosfoproteínfosfatasa separa el grupo fosfato del sustrato proteico *(e)*.

A partir de los trabajos de gran número de neurobiólogos, realizados en el transcurso de los últimos veinte años, hoy en día se sabe que la respuesta de una neurona a la estimulación sináptica consiste en un cambio brusco y pasajero de la permeabilidad –a uno o varios tipos de iones– de la membrana postsináptica. Esta variación de la permeabilidad permite el flujo de iones a través de la membrana, lo que genera una corriente eléctrica y altera, por tanto, el potencial eléctrico de la membrana, es decir, su voltaje. Según cuáles sean los iones que migran y cuál la dirección de su movimiento, el cambio inducido por el neurotransmisor en el potencial eléctrico tendrá una u otra de estas dos acciones opuestas: aumentar o disminuir la probabilidad de que la célula alcance el umbral de excitación necesario para poder generar un impulso nervioso. En este aspecto, la neurona actúa más o menos como un computador e integra los centenares de mensajes químicos inhibidores y excitadores que chocan con su superficie en un momento dado, antes de decidir si producirá o no un impulso.

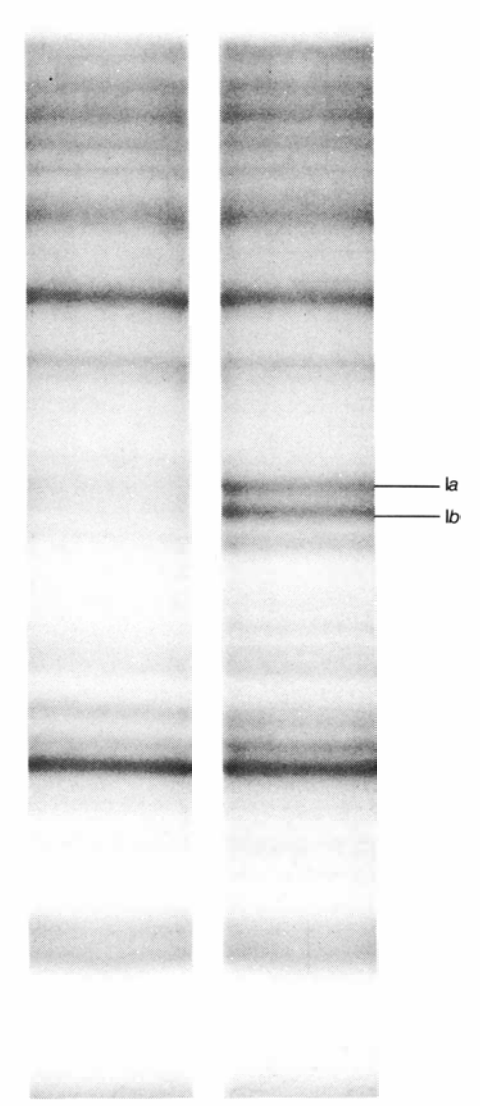
Como nos interesaba estudiar el posible papel del AMP cíclico como mediador de algunos de los cambios que se producen en la permeabilidad de la membrana por acción de neurotransmisores, hace varios años iniciamos una investigación detallada sobre la secuencia de acontecimientos bioquímicos que se dan en el interior de una neurona, tras la unión de un neurotransmisor a su receptor y la activación de la adenilciclasa. Unos experimentos anteriores, llevados a cabo por Edwin G. Krebs y sus colaboradores en la Universidad de Washington, nos habían proporcionado algunas pistas importantes. Al estudiar el mecanismo por el que la adrenalina induce el paso de glucógeno a glucosa en las células musculares, comprobaron que el AMP cíclico, generado en el interior de las células por la adenilciclasa sensible a la hormona, era responsable de que un grupo fosfato procedente de ATP fuera transferido al enzima encargado de la iniciación de la hidrólisis del glucógeno. Luego descubrieron que la transferencia real del grupo fosfato –proceso denominado fosforilación– era llevada a cabo por el enzima proteínkinasa. En la nomenclatura bioquímica, kinasa (o mejor dicho fosfokinasa) es un término reservado para los enzimas que transfieren un grupo fosfato del ATP a otra molécula; en el caso de una proteínkinasa, la otra molécula siempre es una proteína.

Se cree que el AMP cíclico activa a la proteínkinasa, uniéndose a una subunidad inhibidora de la kinasa y provocando un cambio de la forma de la subunidad de tal manera que se disocie del resto del enzima. La pérdida del componente inhibidor potencia la actividad de la kinasa que, inmediatamente, transfiere un grupo fosfato al enzima que controla la vía de degradación del glucógeno. La adición de un grupo fosfato a este enzima desencadena reacciones enzimáticas "en cascada", que se resuelven finalmente en la degradación del glucógeno. Por tanto, la proteínkinasa actúa como lazo de unión entre el AMP cíclico generado por la hormona y la activación de la vía bioquímica que explica los efectos metabólicos de la hormona.

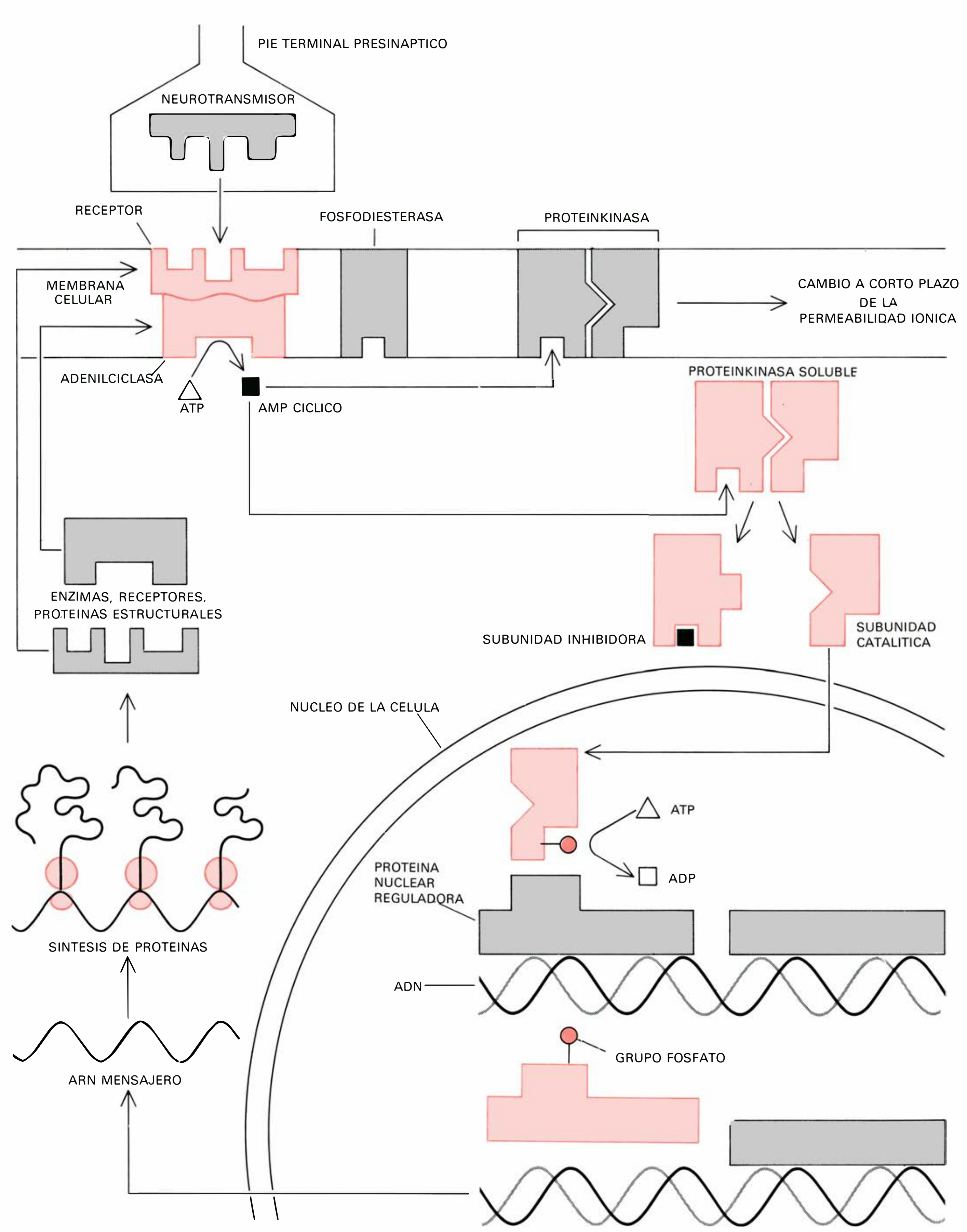
Tomando este esquema como base, empezamos a buscar un sistema análogo en las células cerebrales. Aunque la hidrólisis del glucógeno no es un efecto predominante del AMP cíclico en el tejido nervioso, intuimos que una proteínkinasa similar podría transformar el cambio de los niveles de AMP cíclico provocados por un neurotransmisor en un cambio de permeabilidad a los iones de la membrana postsináptica.

Al cabo de un tiempo relativamente corto, J.-F. Kuo y Eishichi Oiyamoto lograron aislar, en nuestro laboratorio, una proteínkinasa dependiente de AMP cíclico a partir de cerebro de vaca. El enzima se hallaba presente en el tejido cerebral en cantidades relativamente elevadas en comparación con la mayoría de los demás tejidos, y su actividad dependía significativamente de la presencia de AMP cíclico. Lo que es más: la concentración de AMP cíclico necesaria para estimular al enzima hasta alcanzar la mitad de su actividad máxima era similar a la concentración de AMP cíclico que se detecta normalmente en el tejido cerebral. Este resultado indicaba que los incrementos de los niveles de AMP cíclico, provocados por la conocida acción de varios neurotransmisores, podían ser la causa de cambios considerables en la actividad de la proteínkinasa dependiente de AMP cíclico.

Otros experimentos demostraron que la proteínkinasa podía transferir un grupo fosfato a proteínas distintas de las asociadas a la hidrólisis del glucógeno. Hiroo Maeno y Edward Johnson observaron, en nuestro laboratorio, que, cuando se homogeiniza tejido cerebral, las proteínas que más fosforiladas están se encuentran en las fracciones del tejido homogeneizado que contienen una mayor cantidad de fragmentos constitui-



LA FOSFORILACION de las proteínas postsinápticas es una respuesta específica de la membrana sináptica al AMP cíclico. Se incubaron las membranas sinápticas, preparadas a partir de cerebro de rata, en una solución que contenía ATP radiactivo, tanto en presencia como en ausencia de AMP cíclico, con lo que las moléculas de proteínkinasa de la membrana podían transferir un grupo fosfato radiactivo del ATP a sustratos proteicos diversos. Posteriormente, las proteínas fosforiladas se extrajeron de la membrana con un detergente y se separaron unas de otras colocando una pequeña porción de la mezcla en la parte superior de una columna llena de la sustancia gelatinosa poliacrilamida. Al ser sometidas a una corriente de alto voltaje, las proteínas migraron hacia abajo y se separaron en una serie de franjas delgadas, en un orden que dependía de su tamaño. Luego, en el interior de una habitación oscura, se pusieron los geles en contacto con una película sensible a los rayos X: así se podía detectar cuáles eran las franjas que velaban la película, es decir, cuáles eran las que habían incorporado fosfato radiactivo. La reproducción que aquí se ofrece de una película velada pone de manifiesto que, a pesar de que se fosforilaban muchas proteínas por acción de enzimas independientes del AMP cíclico, sólo dos de ellas eran fosforiladas por una proteínkinasa cuya actividad dependiera de la presencia de AMP cíclico. Dichas proteínas se designaron por 1*a* (peso molecular de 86.000 daltons) y 1*b* (80.000 daltons).



LOS EFECTOS A LARGO PLAZO que tiene el AMP cíclico sobre la función de la neurona se producen, hipotéticamente, por activación de una proteínkinasa soluble localizada en el citoplasma. Este enzima, tras ser activado, penetra en el núcleo de la célula donde transfiere un grupo fosfato del ATP a una de las proteínas nucleares reguladoras, que se hallan íntimamente asociadas a porciones del ADN correspondientes a genes inactivos. La fosforilación altera la forma o la capacidad de unión de la proteína reguladora con el ADN, de tal modo que ésta se separa de la doble hélice de ADN, dejando a descubierto la porción de ADN subyacente para permitir la iniciación del proceso de síntesis de proteínas. Se supone que las proteínas cuya síntesis pudiera inducirse de este modo serían: receptores de sustancias neurotransmisoras, adenilciclasas, proteínas estructurales y enzimas encargados de la síntesis o degradación de un neurotransmisor. El AMP cíciclo, generado en los procesos de transmisión sináptica a corto plazo, podría alterar, a largo plazo, las propiedades eléctricas de la célula receptora.

dos por membranas sinápticas, y que dichas fracciones también son las que poseen la concentración de proteínkinasa más elevada.

De este modo se hizo evidente que la membrana postsináptica llevaba todos los elementos esenciales para la transferencia, inducida por el neurotransmisor, de un grupo fosfato a un sustrato proteico situado en la membrana: (1) una adenilciclasa, sensible al neurotransmisor, capaz de generar AMP cíclico como respuesta a un neurotransmisor específico: (2) una proteínkinasa, dependiente de AMP cíclico, capaz de fosforilar un sustrato proteico en presencia de AMP cíclico; y, finalmente (3), el sustrato proteico localizado en la membrana y fosforilado por la proteínkinasa. La transferencia de un grupo fosfato a este tipo de proteína podría alterar de algún modo la permeabilidad de la membrana a los iones, bien directamente cambiando la configuración de la proteína a fin de abrir un poro que permitiera el flujo de los iones a través de la membrana, bien indirectamente como, por ejemplo, afectando a la actividad de una "bomba" enzimática que realizara un transporte físico de los iones a través de dicha membrana.

Una profundización en la investigación puso de manifiesto que las fracciones correspondientes a las membranas sinápticas también contenían la maquinaria enzimática necesaria para detener el proceso de fosforilación dependiente de AMP cíclico y devolver la membrana a su estado de reposo. Este conjunto de enzimas comprende no sólo a la fosfodiesterasa, encargada de degradar al AMP cíclico dando una forma inactiva de AMP, sino también a la fosfoproteínfosfatasa, encargada de separar el grupo fosfato de las proteínas que han sido previamente fosforiladas por acción de la proteínkinasa dependiente de AMP cíclico. Se ha demostrado que varios de los componentes de este sistema enzimático unido a la membrana se encuentran formando un complejo, por lo que así es menor la distancia que deben recorrer los productos de reacción para pasar de un enzima al siguiente y puede ser más rápida la unión o separación del grupo fosfato del sustrato proteico.

A raíz del descubrimiento de que las fracciones de tejido cerebral homogeneizado más ricas en fragmentos de membrana sináptica constituían sustratos excelentes para la proteínkinasa dependiente de AMP cíclico, Tetsufumi Ueda, Bruce K. Krueger y Javier Forn, de nuestro laboratorio, intentaron, junto con Johnson y Maeno, aislar la proteína o las proteínas específicas que estaban siendo sometidas al proceso de fosforilación. Hasta entonces no se había podido averiguar si las proteínas que actuaban de sustrato de la fosforilación representaban un amplio porcentaje de las proteínas cuya presencia en la membrana postsináptica era conocida o tan sólo unas pocas de entre varias docenas identificadas. Ueda y sus colaboradores comprobaron, al incubar fracciones de membrana sináptica en una solución que contenía ATP marcado con radiactividad, por un lado en presencia de AMP cíclico y, por otro, en su ausencia, que sólo se podía observar un aumento notable de la fosforilación en presencia de AMP cíclico en dos o tres de las proteínas que se cuentan por docenas en la fracción de las membranas. Este resultado produjo una gran satisfacción ya que indicaba que los efectos del AMP cíclico sobre la fosforilación de la proteína de la membrana sináptica eran específicos y no implicaban sencillamente una estimulación general del metabolismo celular.

Ultimamente, se ha llevado a cabo en nuestro laboratorio una investigación detallada en torno a una de las proteínas de la membrana sináptica que actúan como sustrato de fosforilación específico de la proteínkinasa dependiente de AMP cíclico: la denominada Proteína I. La Proteína I posee dos subunidades moleculares, I*a* y I*b*, y se halla localizada casi exclusivamente en la región sináptica del tejido nervioso, estando ausente, al parecer, de las organelas celulares de las neuronas (tales como núcleos, mitocondrias y ribosomas), cuya misión no está directamente relacionada con la transmisión sináptica. Además, hasta el momento no se ha detectado la presencia de Proteína I en ninguno de los tejidos no neurales examinados (como, por ejemplo, el corazón y el hígado). Tampoco se encuentra en el cerebro de fetos de rata cuya edad corresponda a etapas de desarrollo anteriores a la formación de conexiones entre neuronas, sino que su primera aparición coincide con la etapa durante la cual se construyen los complejos sinápticos.

La fosforilación de la Proteína I por la proteínkinasa en la membrana sináptica es un proceso sumamente rápido. De hecho, la fosforilación ya se encuentra en el máximo de su funcionamiento al cabo de un período de cinco segundos, que es el tiempo más corto que se haya podido estudiar. Una velocidad elevada en esta reacción es una condición previa para que pueda ser seriamente considerada la posibilidad de que la Proteína I desempeñe algún papel en la generación de los cambios, brevísimos, de permeabilidad de la membrana sináptica a los iones (la duración de los cambios es igual o inferior a varios centenares de milisegundos), que son responsables de la producción de los potenciales postsinápticos.

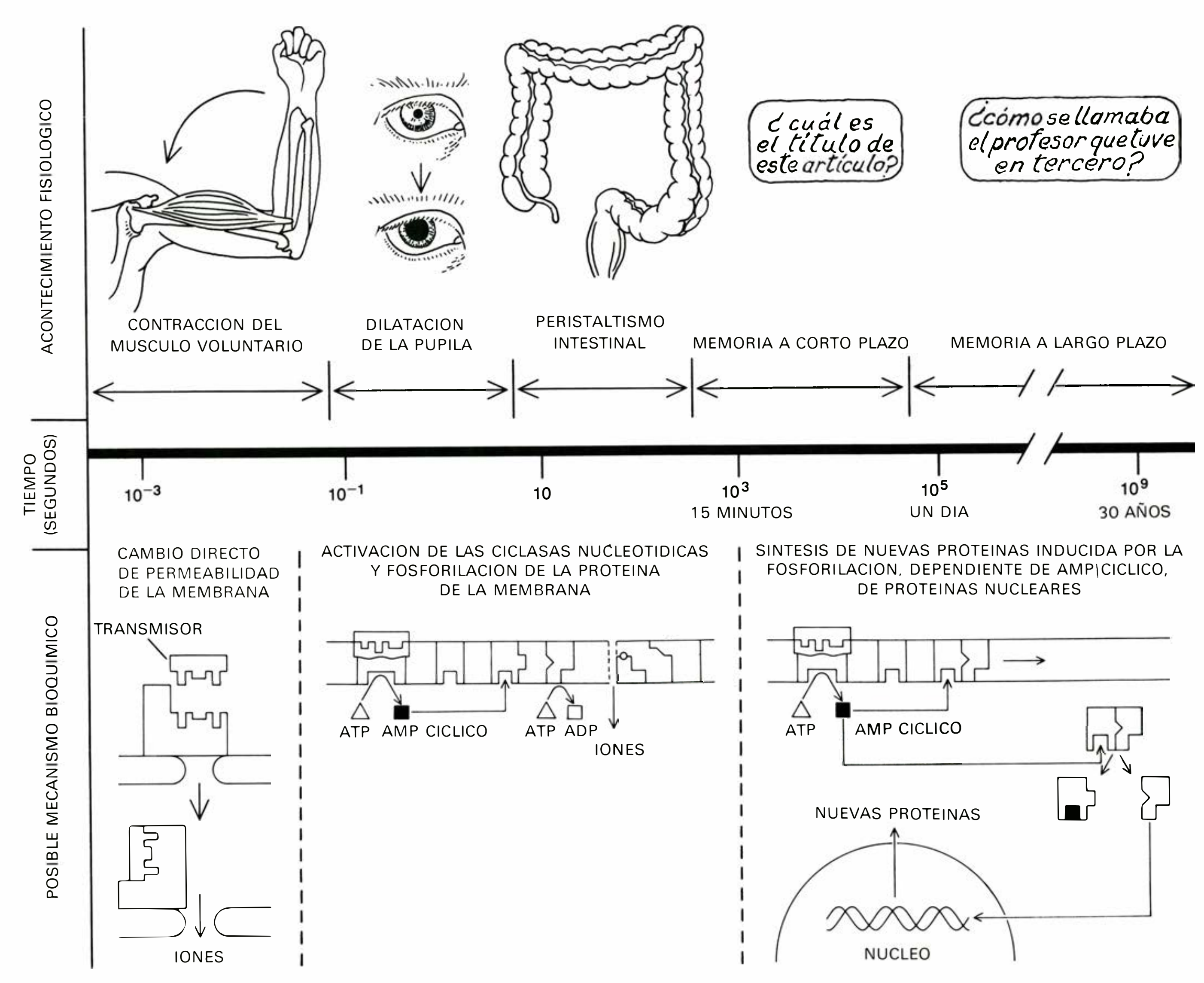
Si la fosforilación proteica dependiente de AMP cíclico es realmente responsable de los cambios de permeabilidad de la membrana sináptica a los iones, a la fuerza debe existir una correlación entre el grado de fosforilación de una determinada o de determinadas proteínas de la membrana y el estado de permeabilidad a los iones en la membrana celular. La corta duración de los cambios de permeabilidad de las membranas pertenecientes al sistema nervioso ha dificultado bastante el establecimiento de este tipo de correlación debido a las limitaciones de la metodología actual. Estas dificultades nos han llevado a estudiar en detalle los mencionados cambios de permeabilidad en otros sistemas no neurales que nos puedan servir de modelo; entre otros, escogimos los glóbulos rojos de pavo y de rana. Dichas células, cuando se encuentran en condiciones adecuadas, responden a las hormonas como la adrenalina con un incremento del flujo de iones sodio y potasio a través de la membrana celular. Parece ser que este incremento, que es debido a un cambio de permeabilidad de la membrana, se produce por mediación del AMP cíclico.

Hace poco, Stephen A. Rudolph comprobó, en nuestro laboratorio, que, cuando a una suspensión de glóbulos rojos se le añade o bien AMP cíclico o bien isoproterenol, un fármaco relacionado con la adrenalina, se fosforila una sola proteína de la membrana, de peso molecular relativamente elevado. Realizó también mediciones simultáneas del grado de fosforilación de esta proteína y del flujo de iones sodio hacia el interior de la célula en presencia de isoproterenol, lo cual le permitió averiguar que el tiempo necesario para que la proteína de la membrana alcanzara la mitad de su máximo grado de fosforilación coincidía estrechamente con el tiempo necesario para que el flujo de iones sodio a través de la membrana alcanzara la mitad de su velocidad máxima.

Los resultados de estos experimentos y de otros complementarios, realizados en varios sistemas distintos, ponen de manifiesto la existencia de una estrecha correlación entre el movimiento de io-

nes a través de las membranas celulares y la fosforilación, dependiente de AMP cíclico, de unas proteínas específicas situadas en dichas membranas. No obstante, cabe resaltar que la correlación no siempre implica una relación causal. En el momento actual, no tenemos pruebas directas de que la fosforilación sea un paso necesario para que se puedan dar los cambios de permeabilidad observados; también podría ser que ambos procesos fueran efectos secundarios de algún otro proceso. La prueba definitiva de una relación causal entre la fosforilación de las proteínas de la membrana y los cambios de permeabilidad sigue siendo uno de los principales objetivos que quedan por alcanzar en la experimentación biológica. Albergamos esperanzas de que, algún día, se podrán separar perfectamente los componentes moleculares individuales responsables del control de la permeabilidad iónica, de tal modo que se podrán colocar en membranas sintéticas con el fin de determinar si la fosforilación previa, dependiente de AMP cíclico, de uno o varios de estos componentes da como resultado cambios de permeabilidad similares a los observados en células intactas.

Debido a la naturaleza compleja de los procesos que implican la mediación del AMP cíclico, los cuales suelen comprender varios pasos bioquímicos, dichos procesos parecen ser más adecuados para regular acontecimientos sinápticos de duración relativamente larga. Este podría ser el motivo por el que las vías nerviosas que utilizan nucleótidos cíclicos son a menudo las que desempeñan un papel de regulación en el sistema nervioso, es decir, las que regulan la actividad nerviosa en vez de iniciarla. Por ejemplo, el AMP cíclico formado en el cerebelo por activación, a nivel del locus coeruleus, de las fibras portadoras de norepinefrina actúa como moderador de la frecuencia "espontánea" de producción de impulsos en las neuronas de Purkinje. En los ganglios del sistema nervioso simpático, las sinapsis en las que la dopamina actúa por mediación del AMP cíclico ajustan el nivel de actividad de otras sinapsis en las que no interviene el AMP cíclico. Incluso



EL CONTINUUM DE LOS ACONTECIMIENTOS SINAPTICOS que se da en el sistema nervioso se extiende desde los que duran tiempos brevísimos hasta otros muy prolongados. En el extremo de tiempos cortos del espectro, tenemos procesos, como la contracción del músculo voluntario, que se inician en el espacio de pocos milisegundos por acción directa del neurotransmisor sobre la membrana, que abre inmediatamente un poro para el paso de iones. Los acontecimientos cuya duración va desde varios centenares de milisegundos hasta minutos, al parecer, se produce de modo indirecto, por mediación de una síntesis de AMP cíclico inducida por el neurotransmisor: el AMP cíclico formado inicia la fosforilación de proteínas de la membrana, lo cual desencadena un cambio relativamente lento de la permeabilidad iónica. Los acontecimientos que duran horas o incluso años, la memoria por ejemplo, quizás impliquen una síntesis de proteínas *de novo* que esté modulada por una fosforilación, estimulada por el AMP cíclico, de las proteínas cuya misión específica consiste en regular la expresión de los genes.

en los casos de trastornos del movimiento, como puede ser la enfermedad de Parkinson, aparentemente el papel del AMP cíclico formado en el núcleo caudado por activación de la adenilciclasa sensible a la dopamina consiste en regular los movimientos, una vez éstos han sido iniciados por señales procedentes de otras áreas del cerebro.

Los cambios de excitabilidad observados en las sinapsis del cerebelo y del ganglio simpático, en las que el AMP cíclico actúa de mediador, pueden tener una duración de varios centenares de milisegundos o incluso mayor, lo que, dentro de la escala de los acontecimientos a nivel de neuronas, representa un tiempo relativamente largo. Unos trabajos recientes, realizados en un ganglio de molusco por Eric R. Kandel, de la Facultad de Medicina y Cirugía de la Universidad de Columbia, y en un ganglio del sistema nervioso simpático de conejo por Benjamin Libet, de la Facultad de Medicina de la Universidad de California en San Francisco, han revelado la existencia de unos acontecimientos que parecen estar relacionados con el AMP cíclico, y cuya duración es incluso superior a los anteriormente mencionados, alcanzando varias horas en algunos casos. Las observaciones de este tipo dan cierto apoyo a la posibilidad de que los acontecimientos sinápticos mediados por nucleótidos cíclicos estén en la base de determinados cambios a largo plazo en el sistema nervioso central humano.

Esta hipótesis se ha ido confirmando por ciertos hechos que indican que la fosforilación proteica dependiente de AMP cíclico podría ejercer una influencia sobre acontecimientos que se dan en el núcleo de la célula. Varios investigadores cuyo pionero fue Thomas A. Langan, de la Facultad de Medicina de la Universidad de Colorado, han demostrado que las histonas, unas proteínas con carga positiva que se unen al armazón de grupos fosfato, con carga negativa, de la doble hélice del ADN, son susceptibles de ser fosforiladas por la proteínkinasa dependiente de AMP cíclico. Cuando las histonas se hallan íntimamente asociadas al ADN, todo parece indicar que inhiben la expresión de la información contenida en el ADN, por estar bloqueándoles el paso a los enzimas encargados de realizar la transcripción del ADN a ARN. La fosforilación por la proteínkinasa hace que la carga de las histonas se vuelva más negativa, debilitando, por tanto, su unión con el ADN.

Desde un punto de vista mecanicista, el vínculo entre la adenilciclasa y la proteínkinasa nos proporciona un sistema elaborado capaz de transformar acontecimientos sinápticos a corto plazo en cambios bioquímicos de mayor duración. No parece descabellado suponer que el AMP cíclico, formado como resultado de la actividad sináptica, podría activar una proteínkinasa que fosforilaría la histona o alguna otra proteína nuclear de función reguladora, provocando así su separación de una porción de ADN correspondiente a genes inactivos. Posteriormente, al ser puestos a descubierto, los genes estarían disponibles para ser transcritos a ARN mensajero y luego traducidos a proteína. De hecho, unos experimentos realizados en el laboratorio de Erimino Costa, del National Institute of Mental Health, sugieren que el AMP cíclico puede inducir la síntesis de la tirosínhidroxilasa, un enzima implicado en la fabricación de dopamina y de norepinefrina.

En consecuencia, el sistema del AMP cíclico podría constituir un mecanismo por el que la activación continuada de una sinapsis determinada conduciría a la síntesis de nuevas moléculas de enzima o de receptor, produciendo así una alteración permanente de las propiedades eléctricas de la neurona receptora. Sería arriesgado afirmar que los cambios mencionados constituyen la base molecular del almacenamiento de información en el sistema nervioso, pero, sin duda alguna, sería una hipótesis muy atractiva. Cabe especular incluso con la posibilidad de que la fosforilación, dependiente de AMP cíclico, de las proteínas de la membrana sináptica fuera un tipo de memoria a corto plazo y que la fosforilación de las proteínas nucleares con la ulterior síntesis de nuevas moléculas proteicas representara un cambio más duradero, es decir, una memoria a largo plazo.

En resumen, en el cerebro existe un continuum de acontecimientos funcionales cuya duración se extiende desde tiempos cortísimos hasta otros mucho más largos. En el extremo de tiempos cortos de la escala, se encuentran los acontecimientos sinápticos, tales como la contracción del músculo voluntario, que duran tan sólo pocos milisegundos; en el otro extremo se sitúan las memorias que pueden durar 50 años o incluso más. Según nuestros conocimientos actuales, parece posible que, a excepción de los acontecimientos más breves, todos dependan en algún modo de mecanismos relacionados con el AMP cíclico.

La mayor parte de este artículo ha puesto el énfasis en el posible papel del AMP cíclico en las sinapsis y en su importancia como mediador de los efectos de muchos neurotransmisores y de los fármacos que afectan a nuestro comportamiento.

La investigación sobre el papel de los nucleótidos cíclicos en el sistema nervioso se ha centrado de forma especial en las áreas anteriormente mencionadas, hecho que en gran parte se debe a las analogías teóricas y operativas entre la acción de las hormonas y la transmisión de los impulsos nerviosos. En cierto modo, se han subestimado otros posibles papeles de los nucleótidos cíclicos en el sistema nervioso, más que nada por una falta de pruebas, en comparación con otros aspectos. La gran cantidad de estudios que se han llevado a cabo en el transcurso de los últimos dos o tres años, tanto en nuestro laboratorio como en otros centros, sugiere, no obstante, que los procesos, en los que actúan como mediadores los nucleótidos cíclicos, están implicados no sólo en el mecanismo de cambios de permeabilidad iónica inducido por neurotransmisores sino también en la regulación de otros fenómenos variados, como pueden ser la activación de enzimas, la síntesis y liberación de neurotransmisores, los movimientos intracelulares, el metabolismo de los carbohidratos en el cerebro y, posiblemente, incluso procesos de crecimiento y desarrollo. No nos extenderemos sobre el tema de estos progresos recientes, pero nos gustaría subrayar su importancia potencial para el conocimiento global del funcionamiento del sistema nervioso.

En concreto, puede ser útil enfocar los efectos del AMP cíclico, sinápticos y no sinápticos, como parte integrante de un sistema conjunto. Podría ser que una estimulación sináptica suficiente no sólo elevara los niveles de AMP cíclico de forma que se alterara la permeabilidad de las membranas sino que también iniciara una secuencia lógica de acontecimientos mediados por los nucleótidos cíclicos. Esta secuencia podría comprender: disminución o incremento de la síntesis de un neurotransmisor como respuesta a la estimulación sináptica, iniciación de movimientos intracelulares destinados a transportar los productos recién sintetizados, activación del metabolismo de los carbohidratos con el fin de colmar las necesidades energéticas que, forzosamente, debe tener la célula, y efectos directos sobre el material genético en el núcleo de la célula, que pueden producir alteraciones del comportamiento a largo plazo, tales como la memoria.

Juegos matemáticos

Claves de nuevo tipo cuyo desciframiento ocuparía unos cuantos millones de años

Martin Gardner

"A pocos podremos convencer de que no es nada fácil inventar un método de escritura secreta que resista investigación cuidadosa. No obstante, puede afirmarse rotundamente que el ingenio humano es incapaz de preparar una clave que el ingenio humano no pueda resolver."

EDGAR ALLAN POE

La tendencia ascendente de las tarifas postales, acompañada del progresivo deterioro de los servicios de correos, podrá o no continuar; pero probablemente esta cuestión carezca de importancia dentro de algunas décadas, al menos en lo que a la mayor parte de las comunicaciones entre particulares se refiere. La razón es simple. A buen seguro, la transferencia de información será mucho más rápida y mucho más barata por "correo electrónico" que por los sistemas postales ordinarios. En breve será posible dirigirse a cualquier teléfono, insertar un mensaje en un dispositivo adecuado y marcar el número de nuestro destinatario. El teléfono del otro extremo imprimirá el mensaje inmediatamente.

Entre las primeras entidades que harán uso sistemático del correo electrónico se encontrarán, presumiblemente, organismos estatales y grandes empresas, a los que seguirán empresas más pequeñas e individuos particulares. Conforme esto vaya sucediendo, será más y más deseable disponer de claves rápidas y eficientes que protejan la información de escuchas electrónicas. Se plantea un problema parecido con la protección de información privada, almacenada en bancos de memoria computerizada, contra fisgones con acceso a la memoria a través de redes de proceso de datos.

Poco puede sorprender que en estos últimos años cierto número de matemáticos se hayan preguntado: ¿Es posible diseñar un método de encifrado que pueda codificarse y decodificarse rápidamente mediante computador, que pueda ser utilizado reiteradamente sin cambiar la clave, y que pueda resistir los más perfeccionados métodos de análisis criptográfico? Sorprendentemente, la respuesta es afirmativa. El descubrimiento se realizó hace apenas dos años, y probablemente supondrá una revolución en toda el área de la comunicación secreta. Es, de hecho, tan revolucionario, que todas las claves anteriores, juntamente con los métodos para descifrarlas, pueden perderse pronto en el olvido.

Un código indescifrable puede ser indescifrable en teoría o indescifrable solamente en la práctica. Edgar Allan Poe, que se consideraba experto criptoanalista, estaba convencido de que no podría inventarse clave alguna que no pudiera también ser "desenredada". Poe andaba equivocado, por supuesto. Durante medio siglo se han estado utilizando claves que no pueden ser descifradas ni siquiera en teoría. Así ocurre con las "libretas de uso único", claves que se usan solamente una vez, para un único mensaje. He aquí un ejemplo sencillo basado en una clave de desplazamiento, llamada a veces clave cesárea porque Julio César la utilizó.

Se escribe en primer lugar el alfabeto, seguido de los dígitos de 0 a 9. (A efectos de codificación, el 0 representará espacios entre dos palabras, y, los otros dígitos, los signos ortográficos.) Se escribe debajo la misma sucesión desplazada cíclicamente a la derecha un número cualquiera de unidades, como se muestra en color en la ilustración de esta página. La clave consiste en tomar cada símbolo del texto llano (el mensaje), buscarlo en la línea superior y sustituirlo por el símbolo situado directamente debajo. Se tiene así una clave de sustitución simple, que fácilmente descifraría cualquier aficionado.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G	H

S	T	U	V	W	X	Y	Z	0	1	2	3	4	5	6	7	8	9
I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Clave cesárea con desplazamiento 10

A pesar de su sencillez, una clave de desplazamiento puede servir de base para una codificación verdaderamente indescifrable. Todo consiste en usar una clave de desplazamiento distinta para cada símbolo del texto llano, dejando que sea el azar quien decida en cada ocasión la magnitud del desplazamiento, lo que se consigue con una ruleta, como la que se muestra en la ilustración superior de la página opuesta. Supongamos que la primera palabra del texto llano sea LOS. Hacemos girar la flecha, y vemos que se detiene en K. Esto nos dice que para cifrar la L hemos de usar una clave cesárea en la que el alfabeto inferior está desplazado 10 pasos a la derecha, con lo que la letra A caerá debajo de la K, como se ve en la figura. En la cifra así obtenida, la L se transforma en una B. El mismo procedimiento se usaría para cada uno de los símbolos del texto. Antes de codificar cada símbolo se gira la flecha, y se desplaza en consecuencia la sucesión inferior. El resultado es un texto cifrado que empieza por B, y una clave de cifrado que comienza por K. Obsérvese que la clave de cifrado será de la misma longitud que el texto llano.

Para usar estas claves de uso único con el fin de enviarle un mensaje a alguien –que llamaremos *Z*–, primero hemos de enviar la clave a *Z*, por ejemplo, mediante un correo de confianza. Después le enviamos a *Z*, quizá por radio, el texto cifrado. Mediante la clave, *Z* lo descifra, y, a continuación, destruye la clave. No debe usarse dos veces la misma clave, porque, si se interceptaran dos de tales mensajes, un criptoanalista podría disponer de estructura suficiente para decodificarlos.

Es fácil comprender por qué las claves de uso único son indescifrables, incluso en teoría. Dado que cualquier símbolo puede venir representado por cualquier otro símbolo, y que cada elección está determinada por el azar, no hay, por así decirlo, estructura interna. Podemos expresarlo de otro modo. Cualquier mensaje que tuviera la misma longitud que el texto cifrado da una decodificación del mensaje tan aceptable como cualquier otra. Aun cuando se llegase a encontrar el texto llano de un mensaje así cifrado, de nada le serviría en el futuro al criptoanalista, porque la próxima vez que se utilice el procedimiento, la clave, elegida al azar, será completamente diferente.

Las "libretas de uso único" se utilizan

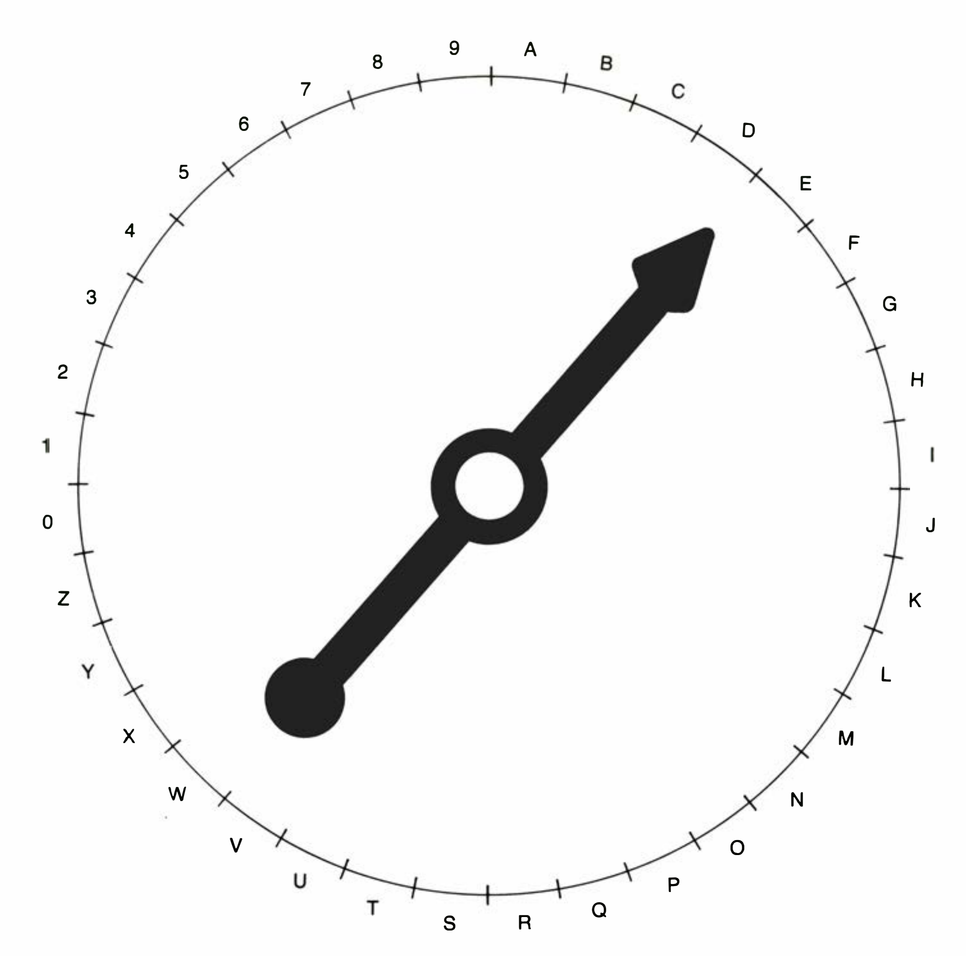
constantemente para transmisión de mensajes entre altos jefes militares, y entre gobiernos y agentes suyos de elevado rango. En realidad, la "libreta" no es sino una larga lista de números aleatorios, quizás impresa en muchas páginas. Tanto el remitente como el destinatario han de disponer de copia de las claves. El remitente usa para cifrar un mensaje la página 1, que después destruye, y así sucesivamente. Cuando el agente ruso Rudolf Abel fue detenido en Nueva York en 1957, tenía una "libreta de claves" en forma de librillo del tamaño aproximado de un sello de correos. David Kahn, que nos cuenta esta historia en su maravillosa narración *The Codebreakers,* explica que el método comúnmente empleado por la Unión Soviética para sus comunicaciones secretas por radio es de libreta de uso único. El famoso "teléfono rojo" entre Washington y Moscú también utiliza este tipo de claves, regularmente suministradas a cada país a través de las respectivas embajadas.

Si la "libreta de uso único" garantiza el secreto absoluto, ¿por qué no utilizarla para todas las comunicaciones secretas? El motivo es que resulta demasiado engorrosa. Cada vez que vaya a emplearse una clave es necesario enviarla por adelantado, y la clave debe ser, cuando menos, tan larga como el mensaje a transmitir. "El problema de producir, registrar, distribuir y eliminar las claves", escribe Kahn, "puede parecer trivial a los profanos en comunicaciones militares, pero en tiempo de guerra el volumen de tráfico llega a desbordar incluso a los equipos encargados de la transmisión de las señales. En un solo día puede ser necesario cifrar centenares de miles de palabras; tan sólo la producción de los millones de caracteres necesarios para las claves sería sumamente costoso y consumiría una enormidad de tiempo. Como cada mensaje ha de tener su propia clave, aplicar el sistema ideal exigiría enviar en cinta, como mínimo absoluto, el equivalente del volumen total de comunicaciones de toda una guerra."

El taxativo dictamen de Poe puede matizarse haciéndolo extensivo solamente a los métodos de codificación que se usan repetidamente y sin cambiar la clave. Hasta muy recientemente, todos los sistemas de codificación pertenecientes a esta categoría se sabían resolubles, al menos teóricamente, siempre que se dispusiera de suficiente texto cifrado y tiempo necesario para realizar su análisis. Así las cosas, se presentó en 1975 un nuevo tipo de claves que alteró radicalmente la situación, ofreciendo una definición de código "indescifrable" procedente de una rama de las ciencias de la computación conocida como teoría de la complejidad. Estas nuevas claves no son absolutamente indescifrables, en el sentido en que lo son las claves de uso único; en la práctica son indescifrables en un sentido mucho más estricto que lo sea cualquier otro código de uso general diseñado anteriormente. En principio, estas nuevas claves pueden ser descifradas, ¡pero sólo tras millones de años de funcionamiento de un ordenador!

Los dos hombres que han provocado este notable avance son Whitfield Diffie y Martin E. Hellman, ingenieros eléctricos ambos de la Universidad de Stanford. Su trabajo fue subvencionado en parte por la National Science Foundation en 1975, y fue publicado en su artículo "New Directions in Cryptography" (*IEEE Transactions on Information Theory,* noviembre, 1976). En dicho artículo, Diffie y Hellman mostraron cómo crear claves indescifrables que no exigen enviar la clave por adelantado y, ni siquiera, que se oculte el método de encifrado. Las claves pueden usarse una y otra vez, pueden cifrarse y descifrarse eficientemente, y con una ventaja extra: el sistema proporciona también una "firma electrónica", que, a diferencia de las firmas escritas, no puede ser falsificada. Si Z recibe un mensaje "firmado" de A, la firma le demuestra a Z que fue A quien verdaderamente envió el mensaje. Además, la firma de A no puede ser falsificada ni por un espía ¡ni tampoco por el propio Z!

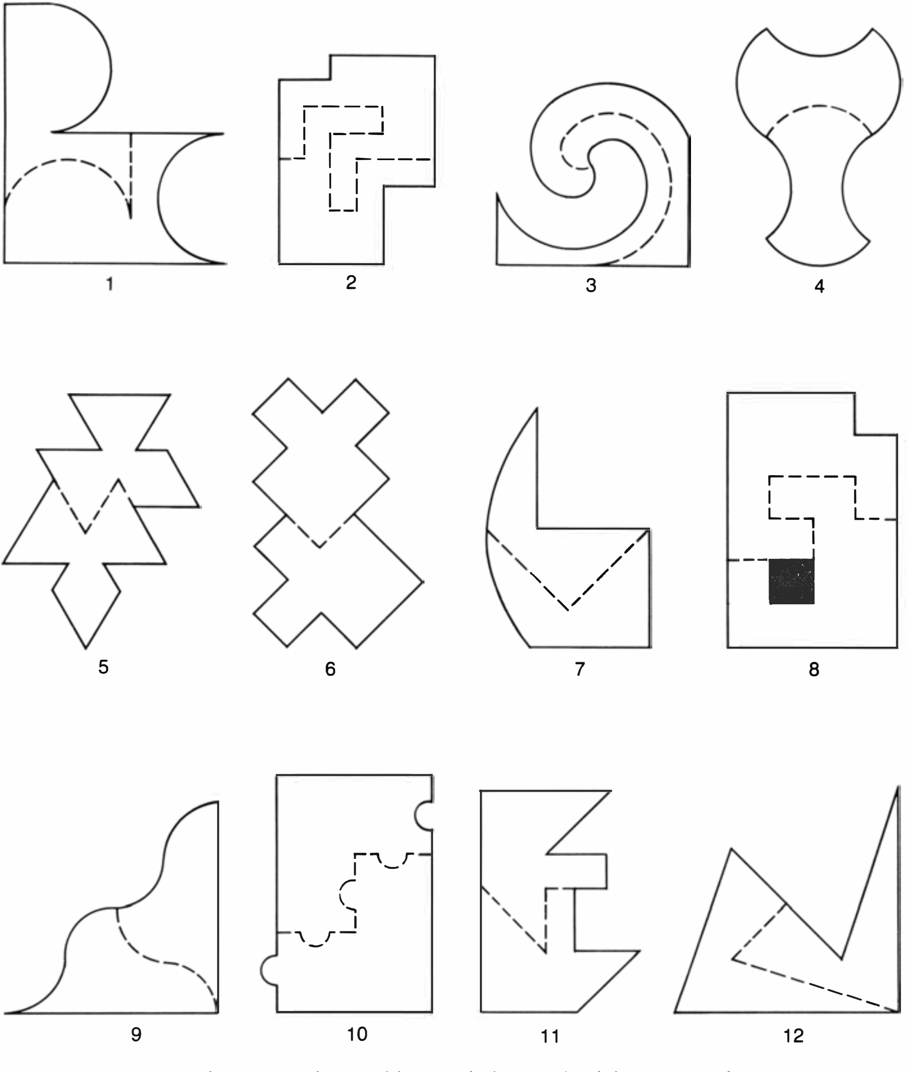
Todas estas características, aparentemente imposibles, son posibles gracias a lo que Diffie y Hellman llaman función trampa de dirección única. Tal fun-



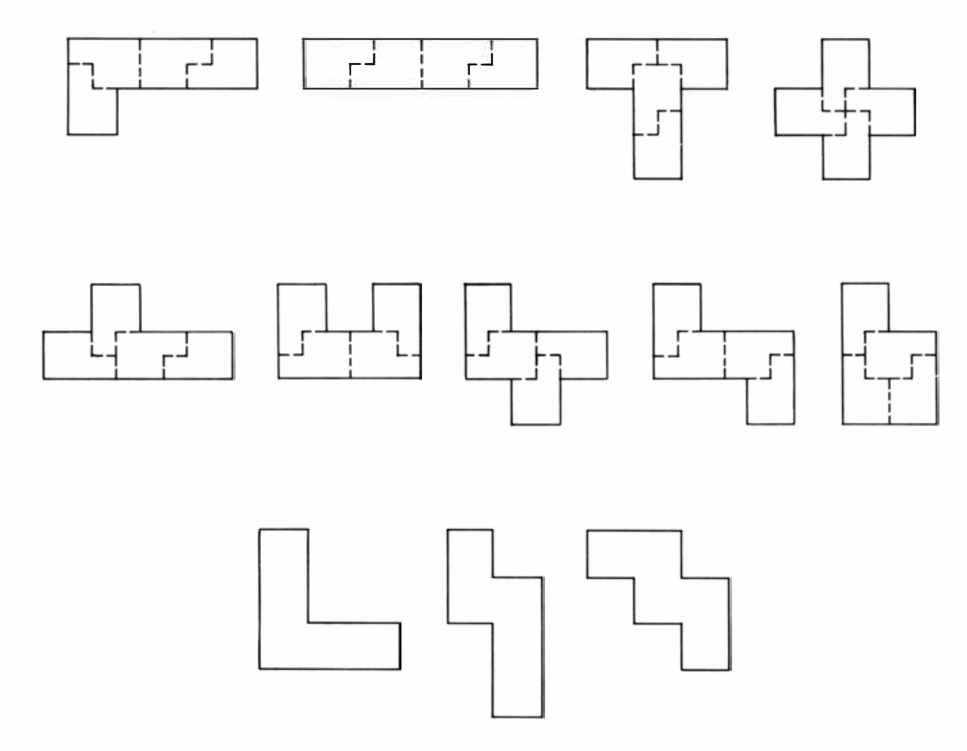
Ruleta para claves de uso único

9686	9613	7546	2206
1477	1409	2225	4355
8829	0575	9991	1245
7431	9874	6951	2093
0816	2982	2514	5708
3569	3147	6622	8839
8962	8013	3919	9055
1829	9451	5781	5154

Texto cifrado de 100 dólares de valor



Soluciones a los problemas de bisección del mes pasado


Descomposición de poliominos en cuatro partes congruentes

ción tiene las siguientes propiedades: (1) transforma todo entero positivo x en un único entero positivo y; (2) tiene función inversa que transforma otra vez y en x; (3) existen algoritmos eficientes para efectuar los cálculos correspondientes tanto a la función directa como a la inversa; (4) si se conocen solamente la función y el algoritmo directos, resulta computacionalmente imposible descubrir el algoritmo inverso.

Esta última es la curiosa propiedad que da nombre a la función. Es como una puerta-trampa: fácil de pasar, pero prácticamente imposible salir por ella; y es imposible salir a menos que se sepa dónde está escondido el botón que la libera. El botón simboliza la "información-trampa". Sin tal conocimiento, no se puede abrir la puerta desde dentro, pero el botón está tan cuidadosamente escondido que la probabilidad de encontrarlo es prácticamente nula.

Antes de dar un ejemplo concreto, veamos cómo tales funciones posibilitan nuevos sistemas criptográficos. Supongamos un grupo de hombres de negocios que desean comunicarse secretos entre sí. Cada uno de ellos prepara su propia función-trampa, juntamente con sus algoritmos directo e inverso. Se publica un manual en el que se expone con detalle el algoritmo de encifrado (directo) de cada compañía. Los algoritmos de decodificación (inversos) permanecen secretos. El manual es de uso público. Cualquiera puede utilizarlo para enviar un mensaje secreto a una de las compañías que figuran en el manual.

Supongamos que no se es miembro del grupo, pero que se desea enviar un mensaje secreto al miembro Z. Primero se transforma el mensaje en un largo número, mediante el procedimiento normal explicado en el manual. A continuación se mira el algoritmo directo de Z, que se usa en el computador propio para cifrar rápidamente el texto. Se le envía a Z este nuevo número. No importa lo más mínimo que el texto cifrado sea aprehendido o interceptado, porque solamente Z conoce el procedimiento secreto de descifrarlo. No hay modo alguno de que un criptógrafo curioso, examinando el algoritmo de codificación, público, de Z pueda descubrir el algoritmo de decodificación correspondiente. En principio, sería posible encontrarlo, pero en la práctica necesitaría un supercomputador y algunos millones de años de trabajo.

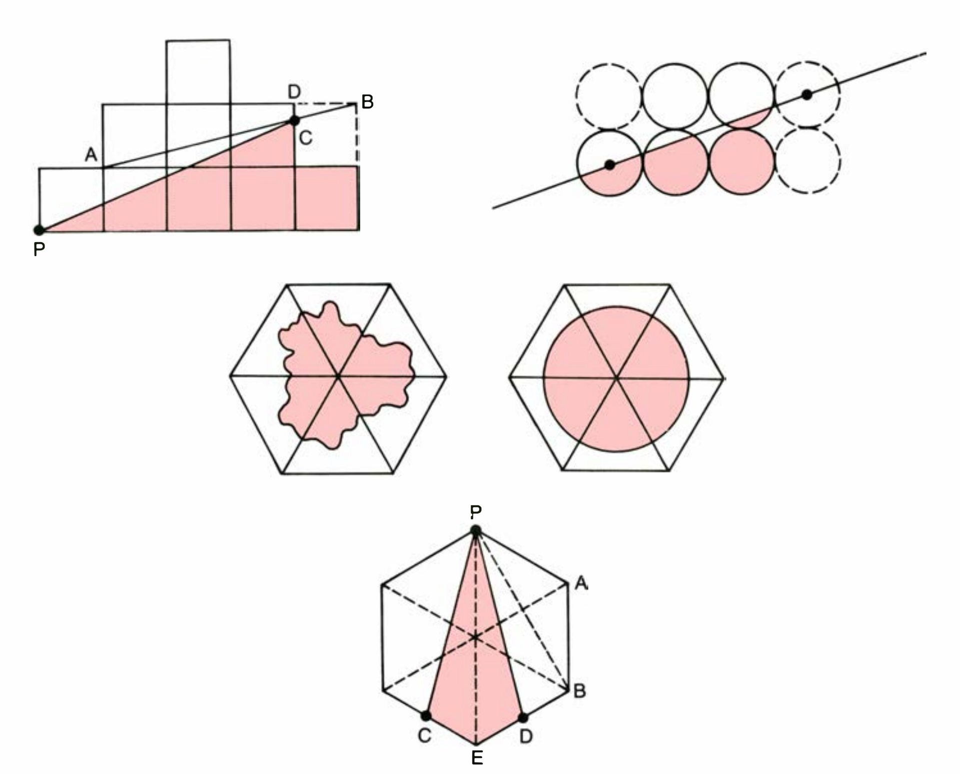
Ningún "extraño" puede firmar un mensaje dirigido a Z, pero sí puede hacerlo cualquier miembro del grupo. La astucia del procedimiento de firma es

diabólica. Supongamos que *A* desea firmar un mensaje dirigido a *Z*. Comienza entonces por codificar su texto llano mediante su algoritmo inverso, secreto. A continuación, vuelve a cifrar el número de texto cifrado antes obtenido, usando esta vez el algoritmo de uso público de Z. Cuando éste recibe el texto así cifrado, lo transforma aplicando primero su algoritmo secreto de decodificación, y, a continuación, aplica el algoritmo de codificación de uso público de *A*. ¡Ya se tiene el mensaje!

Z sabe de este modo que solamente *A* puede haberle dirigido el texto así doblemente cifrado, porque se ha utilizado el algoritmo secreto de *A*. La "firma" de *A* es, evidentemente, imposible de falsificar. *Z* no puede usarla para enviar un mensaje pretendidamente remitido por *A*, porque *Z* sigue sin conocer el algoritmo secreto de decodificación usado por *A*. No solamente eso, sino que de llegar a ser necesario que en algún momento futuro se demostrara ante tercera persona, por ejemplo, un juez en la vista de un pleito, que fue *A* quien realmente envió el mensaje, la demostración sería tan clara que ni *A* ni Z ni nadie podría refutarla.

En su artículo, Diffie y Hellman sugirieron diversas funciones-trampa utilizables en tales sistemas. Ninguna de ellas es exactamente lo que se necesita. No obstante, a comienzos de este año se produjo otro importante avance. Ronald L. Rivest, Adi Shamir y Leonard Adleman, especialistas en ciencias de la computación del Instituto de Tecnología de Massachusetts, pusieron a punto un elegante método para realizar prácticamente el sistema de Diffie-Hellman, utilizando números primos.

Rivest obtuvo su doctorado en ciencias de la computación en la Universidad de Stanford en 1973, y es ahora profesor asociado en el M.I.T. Una vez que dio con la brillante idea de usar números primos para un sistema público de codificación, él y sus colaboradores apenas tuvieron dificultad en encontrar una forma sencilla de conseguirlo. Su trabajo, financiado por becas de la NSF y la Oficina de Investigación Naval, apareció en *On Digital Signatures and Public-Key Cryptosystems* (Memorándum Técnico 82, abril 1977), publicado por el Laboratorio de Ciencias de la Computación, Massachusetts Institute of Technology, 545 Technology Square, Cambridge, Mass. 02139. El Memorándum es gratuito; para recibirlo es suficiente escribirle a Rivest a la dirección indicada, incluyendo sobre para impresos de 24 por 31 cm, convenientemente (0,35 US dól.) franqueado y dirigido al solicitante.

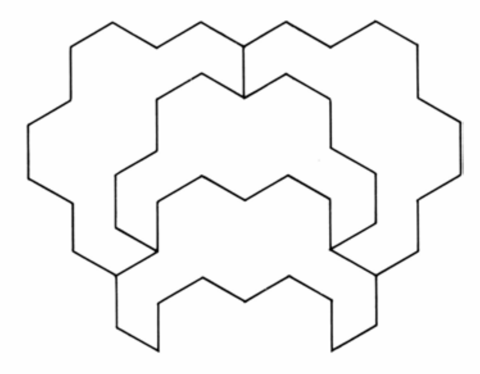


Solución de los cuatro problemas de división en áreas iguales

Para explicar el sistema de Rivest hay que recurrir a la teoría de los números primos. Los programas de computador más rápidos para decidir si un número es primo o compuesto (producto de números primos) se basan en una famosa teoría de Fermat que enuncia que si *p* es primo, y *a* es un número positivo arbitrario menor que *p*, entonces $a^{p-1}=1$ (módulo *p*). Supongamos que se necesite averiguar si un gran número impar *n* es primo (todos los números primos son impares, a excepción del 2). Se elige al azar un número *a* y se eleva a la potencia $n-1$, y a continuación, se divide entre *n*. Si el resto no es igual a 1, *n* no puede ser primo. Por ejemplo, $2^{21-1}=4$ (módulo 21), y, así pues, 21 es número compuesto. ¿Cuál es, sin embargo, la relación entre 2 (el número *a* elegido al azar) y los factores primos de 21, que son 3 y 7? No parece existir relación alguna. Por esta razón, el criterio de Fermat es inútil para hallar números primos. Proporciona, sin embargo, un método rápido para demostrar que un número es compuesto. Además, si un número impar pasa el criterio de Fermat con cierto número de valores aleatoriamente elegidos de *a*, casi con certeza, ese número será primo.

No es éste el lugar de entrar en más detalles sobre programas y criterios algorítmicos para determinar mediante computador el carácter primo de un número –algoritmos extraordinariamente rápidos–, ni para discutir los algoritmos de descomposición factorial de números compuestos –todos ellos de exasperante lentitud–. Me limitaré a unos cuantos hechos, expuestos por Rivest, que ponen de relieve la clamorosa diferencia de tiempo de computador necesario para ambos tipos de prueba. Por ejemplo, para examinar si un número impar de 130 cifras es primo se necesitan como máximo (esto es, cuando el número es primo) unos siete minutos de funcionamiento de un computador PDP-10. Con este mismo algoritmo se tardan solamente 45 segundos en encontrar el primer número primo mayor que 2^{200}. (Se trata de un número de 61 dígitos igual a $2^{200}+235$.)

Contrasta con lo anterior la dificultad de hallar los dos factores primos de un número de 125 o 126 dígitos obtenido multiplicando dos números primos de 63 cifras. Usando el mejor algoritmo conocido y el más rápido de los computadores hoy disponibles, Rivest estima que el tiempo de funcionamiento necesario ¡sería del orden de 40 cuatrillones de años! (Puede verse un buen análisis de los métodos computerizados para descomponer un número en factores primos en *Seminumerical Algorithms*, de Donald E. Knuth, Sección 4.5.4.). Esta imposibilidad práctica de descomponer en factores el producto de dos grandes números primos –al menos durante el futuro previsible– hace viable el sistema de clave propuesto por el M.I.T.



Una tétrada con simetría axial de 18 lados

Para explicar el funcionamiento del sistema, los autores toman como ejemplo de texto llano una paráfrasis de una observación de *Julio César*, de Shakespeare (Acto 1, Escena 2.ª): IT'S ALL GREEK TO ME. (Para mí, todo esto es griego.)

La frase se transforma primero en un único número, mediante la clave normal: A=0,1 B=0,2..., Z=26, usándose 00 para indicar el espacio entre dos palabras. El número resultante es 092019000 1121200071805051100201500l305.

El número completo se codifica elevándolo a una potencia fija s, módulo un cierto número compuesto r. El número compuesto r se obtiene eligiendo al azar (mediante un procedimiento expuesto en el memorándum del M.I.T.) dos números primos, p y q, cada uno de ellos de al menos cuarenta cifras de longitud, y multiplicándolos. El número s ha de ser primo con los números p–1 y q–1. Los números s y r son de conocimiento público, con el fin de ser utilizados en el algoritmo de codificación directa. La operación de codificación puede efectuarse con gran eficiencia, incluso con valores enormes de r; exige menos de un segundo de tiempo de computador.

Los dos factores primos de r, sin embargo, se mantienen en secreto, pues intervienen en el algoritmo inverso, de decodificación, que consiste en elevar el número del texto cifrado a otra potencia t y reducirlo módulo r. Al igual que antes, esta operación requiere menos de un segundo de funcionamiento del computador. Sin embargo, el número t solamente podrá ser calculado por quienes conozcan p y q, los dos números primos secretos.

Si el mensaje es demasiado largo para poder manejarlo como un único número, puede descomponerse en dos o más bloques, y cada bloque ser tratado como un número independiente. No daré ya más detalles, pues son de carácter técnico y están claramente explicados en el memorándum del M.I.T.

Para cifrar IT'S ALL GREEK TO ME, el grupo del M.I.T. ha tomado s=9007 y r=114381625757888867669235779976146 612010218296721242362562561842935706935245733897830597123563958705058989075147599290026879543541.

El número r es producto de un número p de 64 cifras y un número q de 65, seleccionados ambos al azar. El algoritmo de codificación transforma el texto llano (09201...) en el siguiente número de texto cifrado: 19993513149780510045231712274026064742320401705839146310370371740625971608948927504309920962672582675012893554461353823769748026.

Como desafío a los lectores de *Investigación y Ciencia*, el grupo del M.I.T. ha codificado otro mensaje mediante el mismo algoritmo de uso público. El texto cifrado se muestra en la ilustración inferior de la página 97. Su texto llano es una frase inglesa. Primero fue convertida en un número mediante el procedimiento normal antes expuesto, después todo el número fue elevado a la potencia 9007 –ésima (módulo r) mediante el procedimiento abreviado expuesto en el memorándum. El M.I.T. pagará 100 dólares al primero en descifrar este mensaje.

Para demostrar que el mensaje procede verdaderamente del M.I.T. se le ha añadido la siguiente firma: 16717861150380844246015271389168398245436901032358311217835038446929062655448792237114490509578608655662496577974840004057020373.

Para codificar la firma se usó el algoritmo inverso secreto del algoritmo de codificación. Como el lector no tiene algoritmo público de codificación propio, la segunda operación de codificación se ha suprimido. Cualquier lector que tenga acceso a un ordenador y que disponga de las instrucciones del memorándum del M.I.T. puede leer fácilmente la firma aplicando el algoritmo público de codificación del grupo del M.I.T., es decir, elevando el número anterior a 9007 y reduciéndolo después módulo r; resulta 060918192000191512220518002309141900151405000821140418050400041512120l1819. Su traducción (usando otra vez la clave normal) resulta ser FIRST SOLVER WINS ONE HUNDRED DOLLARS. Este texto cifrado solamente pudo ser remitido por el grupo del M.I.T. porque sus componentes son los únicos que conocen el algoritmo inverso utilizado para obtenerlo.

Ni Rivest ni sus asociados tienen pruebas de que en el futuro sea imposible hallar un algoritmo rápido para descomponer en factores números compuestos tan largos como el r que ellos han utilizado, o de que la clave sea resuelta por algún procedimiento en el que no han pensado. Consideran, sin embargo, que tales posibilidades son extremadamente improbables. Evidentemente, todo sistema de encifrado que pueda ser considerado inviolable en el sentido de las libretas de uso único es susceptible de retorcidos análisis, que los modernos criptógrafos, expertos matemáticos, realizan con auxilio de potentes computadores. Si el sistema de codificación del M.I.T. resiste con éxito tales ataques, como parece casi seguro que ocurrirá, el dictamen de Poe será difícil de sostenerse en forma alguna.

Incluso en el caso inverosímil de que el sistema del M.I.T. sea resoluble, existen, con probabilidad, toda clase de funciones trampa capaces de proporcionar claves virtualmente inviolables. Diffie y Hellman están solicitando patentes sobre dispositivos de encifrado basados en funciones trampa que todavía no han revelado. Los computadores y la teoría de complejidad están llevando a la criptografía a una etapa excitante, que quizás adquiera tintes tristes. Repartidos por todo el mundo se encuentran hombres y mujeres de gran inteligencia, algunos de ellos geniales, que han dedicado sus vidas al dominio del análisis criptográfico moderno. A partir de la Segunda Guerra Mundial, incluso aquellas claves gubernamentales y militares que no son de uso único se han hecho tan difíciles de descifrar que el talento de estos expertos va siendo cada vez menos útil. Estas personas se encuentran ahora encima de trampillas a punto de abrirse y sumirlos en las profundidades del olvido.

La ilustración superior de la página 98 muestra cómo pueden descomponerse las doce configuraciones del mes pasado en dos mitades congruentes. La ilustración inferior de la misma página muestra cómo descomponer nueve de los 12 pentominos en las mismas cuatro partes congruentes. Los tres poliominos de orden 5 que permanecen en blanco no pueden descomponerse en cuatro partes congruentes por ningún procedimiento.

La ilustración de la página anterior responde a los cuatro problemas finales de la sección del mes pasado. Para bisecar los nueve cuadrados, dibújese el décimo cuadrado señalado con línea de trazos. Se traza la recta AB a fin de obtener el punto C, y después se une P con C. Suponiendo que los cuadrados tengan lados de longitud 1, CD mediría 1/4 y es fácil comprobar que PC es bisectriz de la figura original. Para la bisección de los cinco círculos se añaden tres círculos

más, como se muestra con línea de trazos. La recta que pasa por los centros de dos círculos descompone en dos, evidentemente, el área total. (Ambos problemas han sido tomados de *A Problem a Day*, de R. M. Lucey, Penguin Books, 1937.)

El hexágono inferior se triseca uniendo *P* con *C* y *D*, puntos medios de dos lados. Supongamos que los triángulos equiláteros tengan área 1. El área de *PAB* es igual a 1, y así, el área de *PBE* es igual a 2; el resto es inmediato. Fui incapaz de encontrar un método de sencillez comparable al anterior para trisecar un pentágono regular con una recta que pase por un vértice.

Los dos hexágonos centrales muestran, como probó Leo Moser, que la curva de longitud mínima que biseca un triángulo es un arco de circunferencia. No importa cuál sea la forma de la curva bisectora, si el triángulo se refleja alrededor de un vértice, como se muestra en la figura, la curva llegará a limitar un polígono curvilíneo cerrado. Dicha figura curvilínea divide en dos partes de igual área al hexágono, y tiene área fija. La figura de mínimo perímetro que abarca una superficie dada es la circunferencia; por consiguiente, las curvas bisectoras de longitud mínima en cada triángulo son arcos de circunferencia. (Este ejercicio está tomado de *Mathematical Quickies*, de Charles W. Trigg, McGraw-Hill, 1967.)

He aquí comentarios a los problemas breves del mes de junio:

Ha sido resuelta la generalización para triángulos de orden *n* del triángulo de bolas de billar, utilizando todos los números consecutivamente a partir de 1. Herbert Taylor descubrió un ingenioso procedimiento para demostrar que es imposible construir ningún TDA (triángulo de diferencias absolutas) con disposiciones triangulares de orden 9 o mayor. Mediante programas de computador se descartaron los TDA de órdenes 6, 7 y 8. Así pues, la única solución de las 15 bolas de billar es el máximo TDA de este tipo.

Solomon W. Golomb propuso tres candidatos para posterior investigación:

1. Si todos los números de un TDA de orden mayor que 5 son distintos pero no consecutivos, ¿cuánto estará obligado a valer el máximo de sus números? (Ejemplo: Es posible construir un triángulo TDA de orden 6 cuyo número máximo sea solamente 22.)

2. Si se usan todos los números consecutivos de 1 hasta *k*, pero permitiendo repeticiones, ¿cuál es el máximo valor de *k* en un TDA de orden *n*? (Ejemplo: Es posible construir un TDA de orden 6 donde *k* tome el valor 20.)

3. ¿Para qué órdenes es posible construir un TDA módulo *m*, siendo *m* el número de elementos del triángulo y siendo los números consecutivos desde 1 hasta *m*? Cada diferencia se expresa módulo *m*. Tales triángulos pueden permutarse de modo que cada elemento bajo la fila superior sea suma (módulo *m*) de los dos números situados sobre él. He aquí, ya permutados, las cuatro soluciones de orden cuatro:

1 6 9 4	2 7 8 3	6 1 4 9	7 2 3 8
7 5 3	9 5 1	7 5 3	9 5 1
2 8	4 6	2 8	4 6
0	0	0	0

Un programa a la inversa preparado por Golomb y Taylor no halló ninguna solución de orden 5; y el Coronel George Sicherman, que inventó el problema de billar primitivo, informa que mediante computador ha descubierto la imposibilidad de soluciones de orden 6. El problema sigue abierto para órdenes superiores.

Robert Ammann, Greg Frederickson y Jean L. Loyer encontraron independientemente una tétrada poligonal de 18 lados con simetría axial mejorando, por consiguiente, la solución de 22 lados que yo había publicado.

Dan Eilers, Allen I. Janis, Scott Kim, P. H. Lyons, Robert Mathews (junto con Martin G. Wallser), James Newton y Mike Tempest encontraron independientemente una segunda solución (ya no hay más) de los recorridos de rey perdido en cuadrados de orden 5.

Cuando terminé mi sección con los aleluyas de longitud decreciente, me referí al de un solo verso como al "último de los cuatro". Draper L. Kauffman, John Little, John McKay, Thomas D. Nehrer y James C. Vibber se encontraron entre los primeros que me indicaron que debí llamarlo "penúltimo de los cinco". El quinto, evidentemente, no tiene ningún verso.

Tom Wright, de Ganges, Columbia Británica, me escribió: "Estuve muy interesado con la paradoja versificada, especialmente con los aleluyas decrecientes de dos versos y un verso. Me pregunté, en efecto, si había Ud. incluido la aleluya sin verso alguno (referente al hombre de Nepal), y estuve observando cuidadosamente para ver si faltaba. Durante el examen, fue mi primer impulso suponer que efectivamente faltaba, pues no se le había dejado sitio, pero un segundo análisis me hizo ver que un poema sin versos, al no necesitar espacio, podría muy bien encontrarse allí. Incapaz de resolver la paradoja por método lógico alguno, me veo en la precisión de preguntarle humildemente si se imprimió o no un aleluya sin versos en el nulo espacio a ellos no reservado".

Un texto en clave que venció a Poe

"Ge Jeasgdxv,

Zij gl mw, laam, xzy zmlwhfzek ejlvdxw kwke tx lbr atgh lbmx aanu bai Vsmukkss pwn vlwk agh gnumk wdlnzweg jnbxvv oaeg enwb zwmgy mo mlw wnbx mw al pnfdcfpkh wzkex hssf xkiyahul. Mk num yexdm wbxy sbc hv wyx Phwkgnamcuk?"

En 1839, en una sección que regularmente escribía para un semanario de Filadelfia, *Alexander's Weekly Messenger*, Edgar Allan Poe desafió a los lectores a que le enviasen criptogramas (claves de sustitución monoalfabética), afirmando que los resolvería "en el acto". Un lector llamado G. W. Kulp le envió un texto cifrado, escrito a mano, en cursiva. Se imprimió como arriba se muestra en el número correspondiente al 26 de febrero de 1840. Poe "demostró" en la correspondiente sección que la clave era un chasco, "una jerga de caracteres tomados al azar, sin significado ninguno".

En 1975, Brian J. Winkel, matemático del Albion College, y Mark Lyster, químico de la clase de criptología de Winkel, descifraron el texto de Kulp. No se trata de una simple sustitución –en esto Poe estaba en lo cierto–, pero tampoco es un texto absurdo. Difícilmente podría Poe ser criticado por sus duros juicios. Además de un grave error de Kulp, hay quince errores de menor importancia, probablemente errores del tipógrafo al leer el manuscrito en cursiva.

Winkel es editor de una nueva revista cuatrimestral, *Cryptologia*, que se puede adquirir en el Albion College, Albion, Mich. 49224, por 16 dólares anuales. En la revista se hace hincapié en los aspectos matemáticos y computacionales de la criptología. En el primer número (enero de 1977) se cuenta la historia del texto cifrado por Kulp, que se presenta como desafío a los lectores. Hasta ahora, solamente tres lectores han podido resolverlo. Daré su solución el mes próximo.

Taller y laboratorio

Un inclinómetro sensible de mercurio que puede actuar como sismómetro

C. L. Stong

Nadie puede asegurar de antemano cuándo va a producirse un terremoto. Sin embargo, investigadores japoneses han observado que, en algunos casos, la superficie de la tierra puede inclinarse muy ligeramente, a lo largo de una línea de falla, un poco antes de que se produzca un seísmo. Los sismólogos han explotado la idea ensayando un instrumento llamado inclinómetro. La parte esencial de un tipo de tales aparatos consiste en dos recipientes cilíndricos que contienen mercurio, algo separados entre sí y que están colocados sobre una base horizontal rígida. Los dos recipientes se hallan interconectados por su base mediante un tubo metálico. Un tubo análogo, pero de plástico, une las partes superiores de los recipientes, en donde hay aire.

Cuando el conjunto se somete a un balanceo, en dirección apropiada, el mercurio pasa de un recipiente a otro. El cambio relativo de nivel del mercurio se mide ininterrumpidamente mediante un circuito eléctrico en el cual las dos superficies libres del metal actúan de placas móviles de un condensador eléctrico diferencial. Una versión rudimentaria del instrumento puede detectar una inclinación de sólo 0,002 segundos de arco, que es aproximadamente el ángulo bajo el cual aparecería un objeto de un milímetro de altura visto desde una distancia de 100 kilómetros.

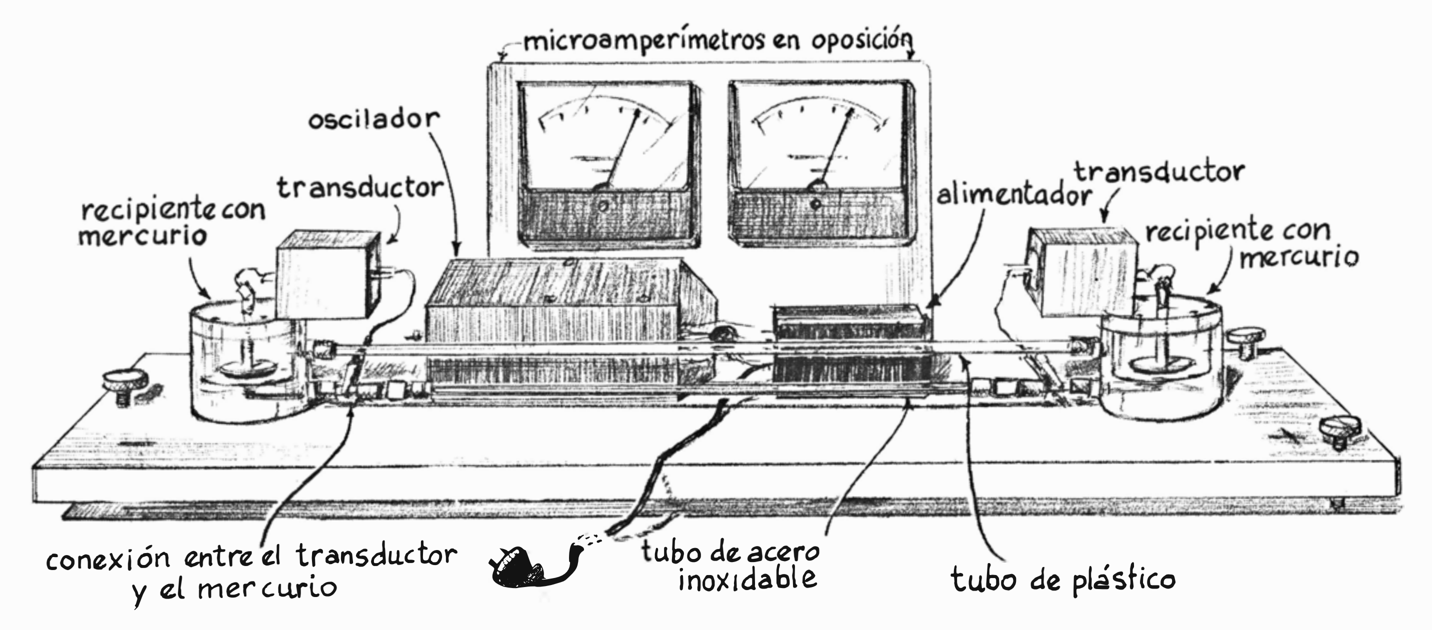
Un inclinómetro de esta clase puede funcionar como un sismómetro barato. El dispositivo es particularmente sensible a las ondas sísmicas que vibran en la dirección de los tubos que unen ambos recipientes. Las aceleraciones experimentadas motivan que el mercurio vaya de un recipiente a otro; dicho líquido actúa de modo análogo al péndulo de un sismómetro corriente, pero el dispositivo es mucho más estable y más fácil de ajustar. Vale la pena reproducir las instrucciones para construir un instrumento de este tipo, con alguno de los efectos sísmicos que puede medir, redactadas por William W. Gile, ingeniero del Laboratorio Sismológico del Instituto de Tecnología de California:

"En la conversación diaria entre la gente del sur de California se habla con mayor frecuencia de la falla de San Andrés que del tiempo, aunque no podamos controlar ninguna de las dos cuestiones. En el lado marino de dicha falla la corteza terrestre se desliza hacia el norte unos cinco centímetros por año. A consecuencia de lo cual, las rocas subyacentes resultan flexadas, se producen desgarros y, finalmente, la falla de San Andrés tiembla.

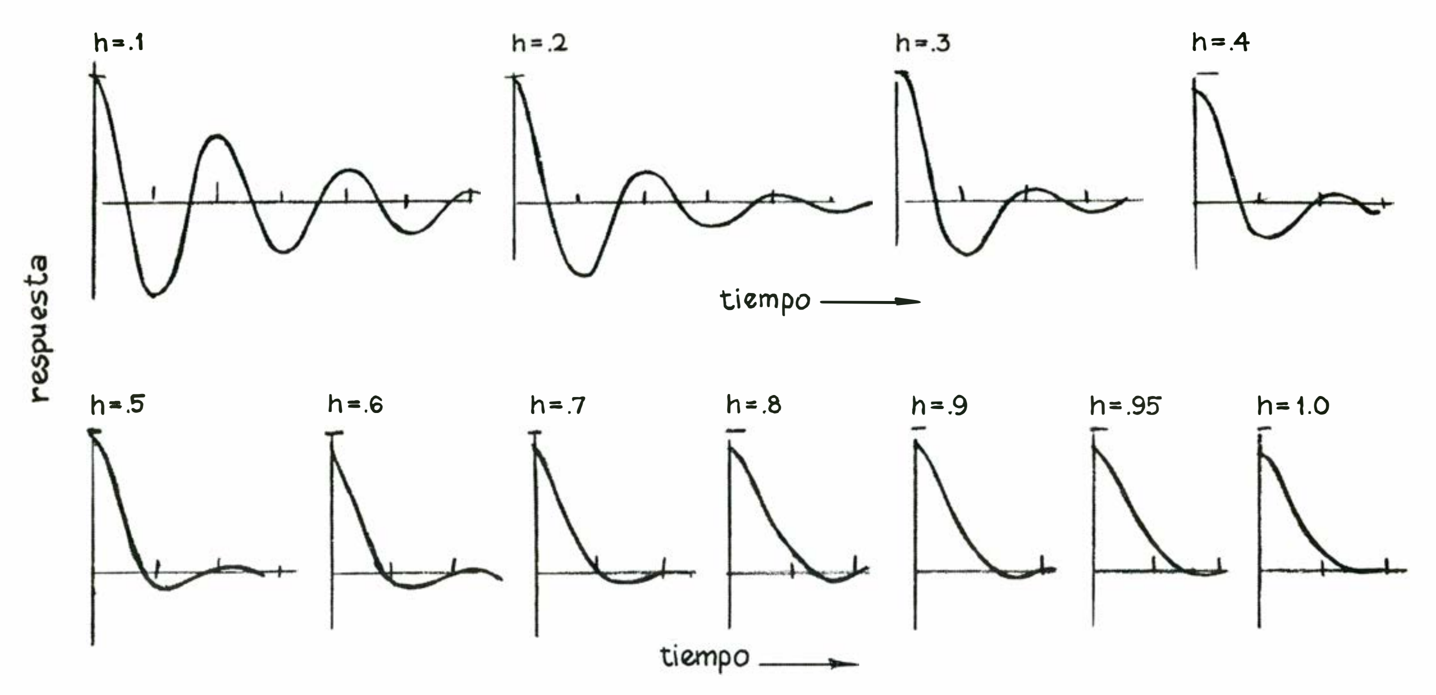
"El reajuste, expresado en términos de centímetros, es impresionante. Así, en 1857 la cerca circular, de unos 6 metros de diámetro, de un corral estaba sobre la falla. El 8 de enero de ese mismo año las tensiones acumuladas a ambos lados de la grieta alcanzaron el valor crítico y se presentaron fracturas de rocas, y cuando los escombros se reposaron la cerca tenía forma de una S. Durante los 116 años siguientes se han ido acumulando las tensiones y, si el movimiento relativo a lo largo de la grieta ha sido de cinco centímetros por año, las placas rocosas habrán experimentado una flexión de cerca de seis metros.

"Por tales razones la predicción de los terremotos se ha convertido en un foco de atención en diferentes ramas de la geofísica y de la sismología. Son varias las técnicas predictoras en las que se tiene depositada la confianza. Pero hay que señalar que aunque no se dispone de un método para predecir los terremotos, las investigaciones actuales no excluyen la posibilidad de lograrlo. Esta esperanza real, aunque débil, mantiene el entusiasmo de los dedicados a trabajar en el campo de la instrumentación sísmica.

"El esfuerzo ha sido premiado en los últimos años con el desarrollo de algunos aparatos notablemente sensibles. Por ejemplo, actualmente, existen medidores de tensiones que de una manera rutinaria aprecian fuerzas equivalentes a las que se desarrollarían si una placa rocosa de la



Inclinosismómetro diseñado por un grupo del Instituto de Tecnología de California



Ejemplos de oscilaciones con diversos valores del factor de amortiguamiento **h**

anchura de los Estados Unidos aumentara en el espesor de un cabello humano. Asimismo existen sensores de temperatura que se emplean en perforaciones profundas que pueden detectar variaciones térmicas, a través de las líneas de falla, que aprecian milésimas de grado centígrado. Finalmente, indicaremos que están disponibles inclinómetros de líneas de base largas que pueden detectar cambios angulares del orden de 10^{-12} radianes, que es aproximadamente el ángulo que corresponde a un segmento de medio milímetro colocado en la luna y visto desde la tierra.

"Algunos de los dispositivos son relativamente baratos y fáciles de ajustar. Un ejemplo lo constituye el inclinómetro de base corta *(véase la ilustración de la página opuesta).* Hemos empleado versiones del mismo en las islas Hawaii para medir cambios en la inclinación de la corteza terrestre ocasionados por las erupciones del volcán Halemaumau. El instrumento trabaja también como sismómetro sensible. Combinado con un registrador de plumilla adecuado, puede revelar las mareas terrestres: es decir, el movimiento diario de ascenso y descenso de la corteza terrestre provocado por las fuerzas de marea ejercidas por la luna y el sol. Su bajo coste, la facilidad de construcción y su alta estabilidad y buena sensibilidad determinan que sea el instrumento ideal para aficionados y profesionales.

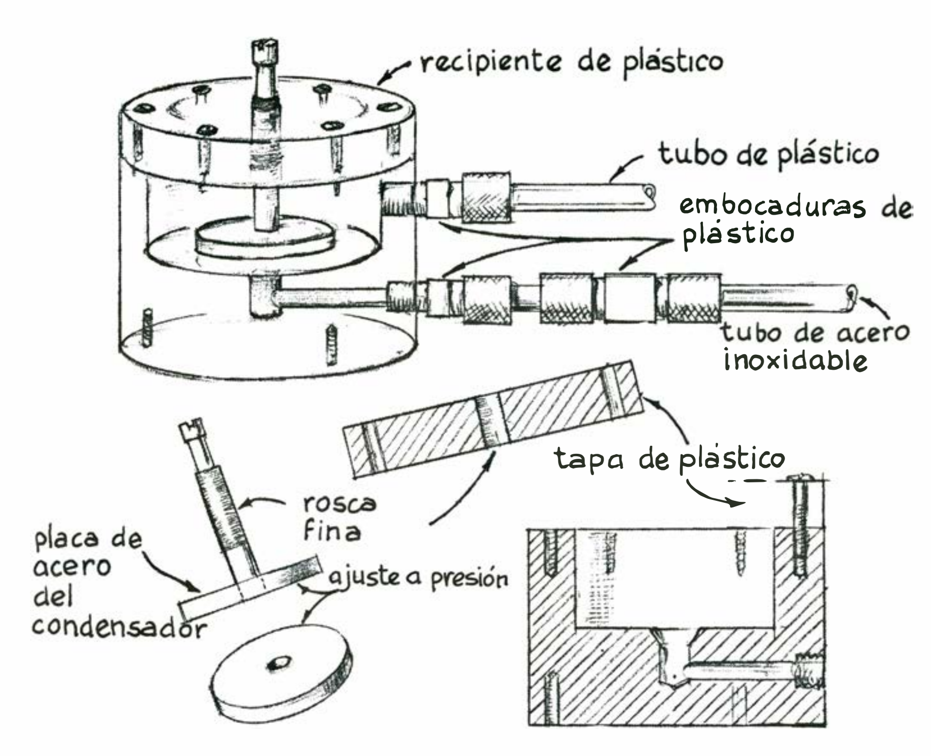
"Los sismómetros convencionales a base de un péndulo dependen de ajustes críticos, según la perturbación sísmica, que determinan el ritmo del péndulo, el número de oscilaciones que realizan antes de pararse y la posición en la que permanecen en reposo. Aunque el mercurio del inclinómetro actúa como un péndulo no necesita de tales molestos ajustes. El período de su oscilación y el amortiguamiento (la rapidez con la cual se para, después de experimentar una perturbación) vienen determinados por las propiedades físicas del mercurio, junto con las dimensiones de los recipientes y de los tubos de conexión. Tales características no cambian de modo apreciable en el transcurso del tiempo. Por ello el instrumento es estable. Una regla normal en la instrumentación sísmica expresa que el período del péndulo –intervalo necesario para completar una oscilación en movimiento de ida y vuelta– debe ser mayor que el de la onda sísmica más larga que haya que registrar. Esta condición es necesaria si se quiere medir el período de las vibraciones sísmicas. En efecto, la masa pendular permanecerá en reposo todavía aunque su soporte vibre, si el período del péndulo es largo respecto al de la vibración. En otras palabras, la masa estacionaria del péndulo sirve de referencia en la medición del movimiento relativo de la corteza terrestre. El principio se aclara mediante un interesante experimento.

"Se toma un cordel de unos 30 cm de longitud, y por uno de sus extremos se ata a un pequeño peso. Por el otro extremo, se sujeta el cordel con la mano, de modo que la parte de la que cuelga el peso sea de 25 cm. Se mueve horizontalmente la mano a un ritmo de cinco oscilaciones por segundo. Se observa que el peso apenas se mueve. Nuestra mano representa la tierra y, el peso, la lenteja del péndulo del sismógrafo; así puede observarse el movimiento relativo terrestre.

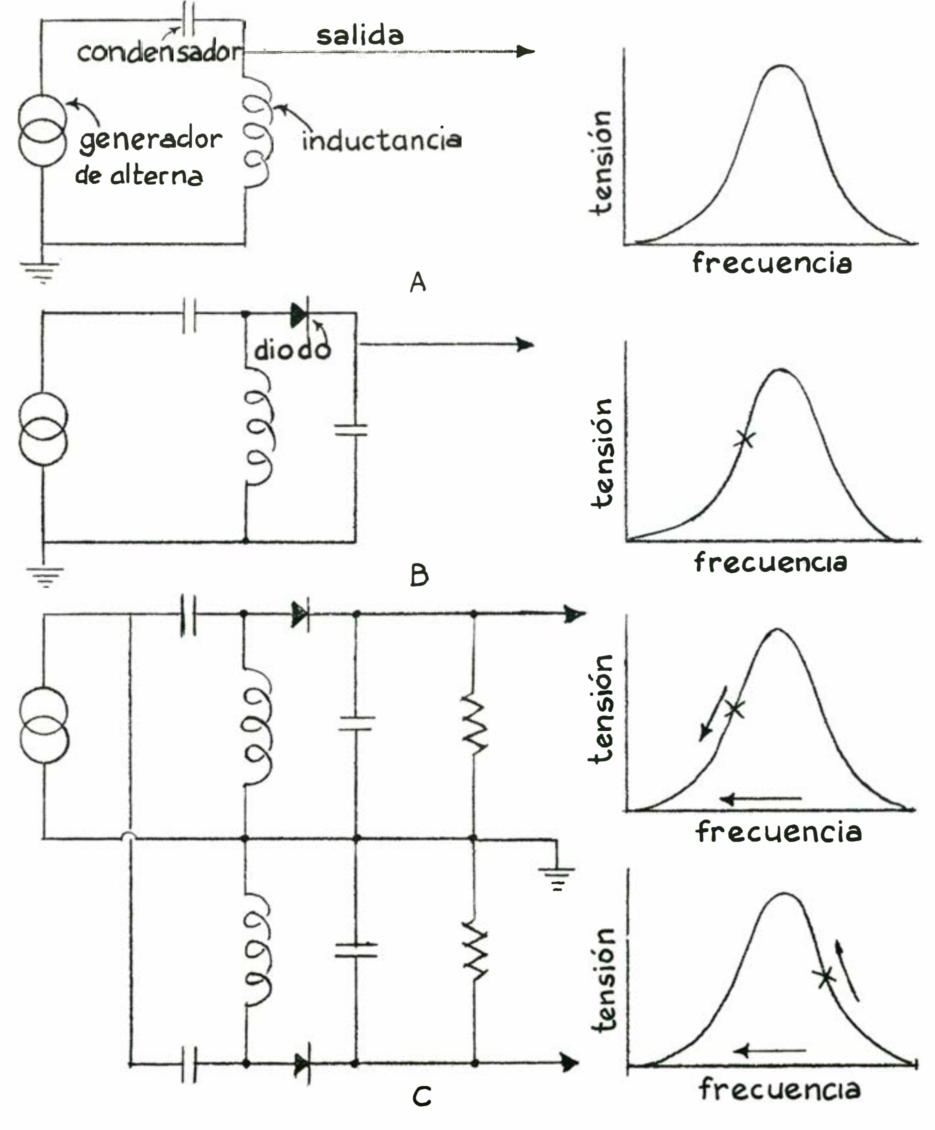
"Repítase la experiencia cambiando el ritmo del movimiento de nuestra mano de modo que sea sólo de una oscilación por segundo. Se observa, en seguida, que el peso oscila con violencia de un lado para otro. La causa reside en que un segundo coincide aproximadamente con la duración del período natural con el que vibra un péndulo de 25 cm, esto es, una oscilación por segundo es la frecuencia propia de la vibración para la cual se presenta resonancia. Si el período de un sismógrafo coincidiera con el de una vibración terrestre se tendría una medida muy exagerada del seísmo correspondiente.

"Párese la mano. El péndulo sigue moviéndose pero con amplitud decreciente. Se dice que está poco amortiguado. El efecto de amortiguamiento resulta de la supresión del movimiento del péndulo por pérdida de energía en el rozamiento. Se dice que un péndulo está críticamente amortiguado cuando se para tras cada vibración. Los buenos sismógrafos que emplean péndulos mecánicos llevan dispositivos ajustables para disipar la energía inadecuada.

"Finalmente, muévase la mano a uno y otro lado lentamente con un ritmo correspondiente a 10 segundos para una oscilación completa. Se observará que el movimiento del peso sigue el movimiento de nuestra mano. Si el ajuste de un péndulo correspondiera a tal situación no apreciaría las sacudidas sísmicas



Detalle de uno de los recipientes de mercurio



Evolución del circuito básico discriminador

de largo período. Para precisar, en el caso de un sismógrafo de desplazamiento, tal como el ilustrado con el experimento, la sensibilidad disminuye con un ritmo de 12 decibelios por octava para vibraciones sísmicas de frecuencia menor que la que corresponde a las vibraciones libres del péndulo. Esto es, cada vez que se duplica el período del seísmo, la amplitud del movimiento del péndulo se reduce a la cuarta parte.

"La tierra muestra un período propio de 53 minutos por ciclo. Es la frecuencia con la que la tierra "suena como una campana" para un gran terremoto. Un péndulo ordinario de este período debería tener una longitud superior a 2400 kilómetros. El período, expresado en segundos, se calcula con aproximación mediante la fórmula $6{,}28\sqrt{l}$; en donde l es la longitud del péndulo expresado en decenas de metro.

"Como ya hemos mencionado, el mercurio de un sismómetro de base corta se comporta como un péndulo. Su período propio está determinado por el diámetro interior de los recipientes y de los tubos de conexión así como por la distancia entre los recipientes. Vale la fórmula siguiente: $T = 2\pi(D_1/D_2)\sqrt{L/2g}$; en donde T representa el período expresado en segundos; D_1 y D_2 son, respectivamente, los diámetros interiores de los recipientes y de los tubos de conexión expresados ambos en la misma unidad, por ejemplo, centímetros; L es la distancia entre centros de los dos recipientes, que se expresará en centímetros, si se toma como valor g de la aceleración de la gravedad 980 cm/s².

"Un instrumento de tamaño adecuado y excelentes características, que es muy apropiado para ser construido por un aficionado, muestra una distancia entre centros de los recipientes de 60 cm con un diámetro interior de cinco centímetros. Los recipientes deben conectarse mediante tubos de acero inoxidable de unos 2'5 mm de diámetro interior. El período de este sismómetro que recomendamos se obtiene aplicando la fórmula que hemos dado. Resulta: $T = 2\pi(5/0{,}25)\sqrt{60/2.980} = 22$ segundos.

"¿Por qué no se usan recipientes más anchos en relación a los tubos de conexión a fin de obtener períodos de una hora o más? En teoría, el período puede aumentarse tanto como se quiera. La práctica indica que el mercurio se mueve menos en los recipientes cuando aumenta la razón de diámetros. La sensibilidad del instrumento para las ondas largas no aumenta con el período.

"Es instructivo conocer el método de cálculo del factor de amortigua-

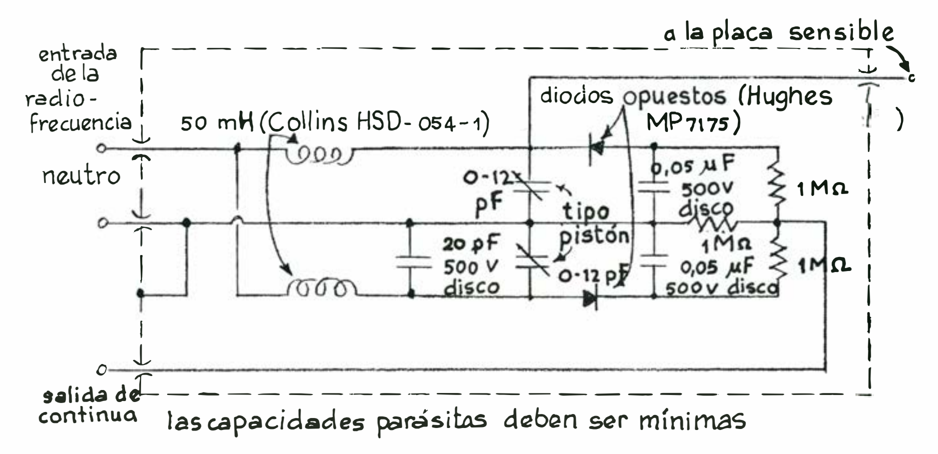
miento del péndulo. Como ya hemos indicado, con el nombre de amortiguamiento se expresa la tendencia de un péndulo a vibrar con amplitud decreciente después de haber sido puesto en movimiento. Dicha tendencia se traduce mediante un factor h que se aproxima a cero en el caso de un péndulo muy poco amortiguado. En el otro caso extremo vale uno cuando el péndulo vuelve a su posición de equilibrio y permanece en ella exactamente, al cabo de medio período.

"En el caso del péndulo de mercurio, el factor de amortiguamiento h se calcula mediante la fórmula $2uT/[\pi d(D_2/2)^2]$, en donde $u=0{,}01554$ (la llamada viscosidad del mercurio expresada en la unidad poise), T es el período del péndulo en segundos, $d=13{,}55$ (la densidad del mercurio en gramos por centímetro cúbico) y D_2 es el diámetro interior de los tubos de conexión expresado en centímetros. El factor de amortiguamiento del sismómetro que hemos recomendado se obtiene sustituyendo los valores adecuados en la fórmula precedente. Resulta $h=2 \times 0{,}01554 \times 22/[3{,}1416 \times 13{,}55 \times (0{,}25/2)^2]=1{,}03$. El péndulo está algo sobreamortiguado, es decir, se para en menos tiempo que el necesario para completar un ciclo de oscilación.

"Las proporciones pueden variarse con objeto de cambiar tanto el período como el factor de amortiguamiento. En la mayoría de instrumentos profesionales el factor de amortiguamiento está comprendido entre 0,7 y 1. Los cálculos indican que se consigue un período de 25 segundos y un factor de amortiguamiento de 0,76 conectando dos recipientes de 5 centímetros, con una distancia entre centros de 1,2 metros, mediante tubos de acero inoxidable de 3,1 milímetros de diámetro interior. Los experimentadores pueden modificar las dimensiones para conseguir otras características.

"Los recipientes cilíndricos pueden ser de plástico transparente. Si el experimentador no dispone de torno, los recipientes pueden fabricarse con una base plana y con tubo de pared gruesa. Los recipientes deben ser cilíndricos, y no rectangulares, para evitar el desarrollo de turbulencia no lineal. Esta no queda eliminada de modo diferencial y se añade al "ruido" de sistema.

"Los dos recipientes se conectan entre sí por medio de tubos de acero inoxidable. Se ha elegido dicho material pues proporciona un contacto eléctrico con el mercurio y así se evita la formación de amalgama en el punto de unión con el circuito eléctrico. El segundo tubo, que está situado por encima de la superficie libre del mercurio, sirve para igualar la presión en los recipientes. Puede ser de plástico.



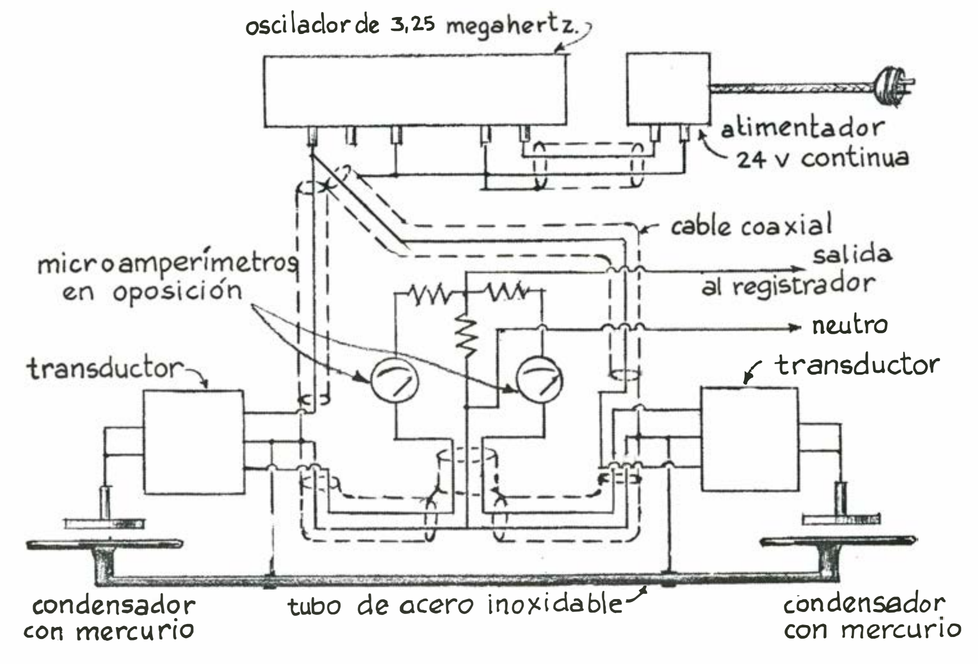
Esquema del circuito del discriminador inverso

"Conviene llenar los recipientes de los aparatos de base larga con nitrógeno para evitar la oxidación en las superficies libres del mercurio. A veces, en tales instrumentos se emplean líneas de base de 30 metros. Los tubos son de plástico de butirato, pues el coeficiente de dilatación de tal material es aproximadamente igual al del mercurio. Los contactos eléctricos con el metal están asegurados por las embocaduras de acero inoxidable que acoplan el tubo de plástico a los recipientes.

"Se mantiene un electrodo fijo en forma de disco pulido de acero inoxidable por encima de la superficie del mercurio en cada recipiente, mediante una varilla roscada que pasa a través del centro de la tapa. Estos electrodos constituyen las placas fijas sensoras del condensador del péndulo de mercurio. Deben formar un ángulo recto con la varilla soporte.

"La varilla actúa como tornillo para ajustar la distancia entre la superficie del electrodo y la del mercurio. El paso de rosca debe ser lo más pequeño posible. Hemos empleado el de 0,3 milímetros. Si no se dispone de útiles apropiados los roscados se encargan en un taller del ramo.

"Si la tapa de plástico tiene un grosor de un centímetro o poco más y se hace un agujero ligeramente más estrecho que la varilla, puede no ser necesario el roscar, pues la varilla roscada dejará una débil huella helicoidal en el plástico. La base para sostener el instrumento debe ir provista de tornillos de nivelación. Puede improvisarse con el perfil de aluminio, en forma de tubo rectangular, frecuentemente empleado en marcos de puertas para oficinas. Fácilmente pueden conse-



Esquema del cableado del inclinosismómetro

guirse trozos de longitud adecuada de dicho material a bajo coste.

"Un transductor electrónico, denominado discriminador inverso, convierte el desplazamiento experimentado por el mercurio en una señal variable de potencial eléctrico. La tensión correspondiente puede actuar sobre un registrador de plumilla. En esencia, el transductor consta de un circuito semejante a los empleados en los aparatos de radio para detectar señales de frecuencia modulada. Tales circuitos reciben el nombre de discriminadores. Suelen incluir un condensador fijo y una bobina que ayuda a detectar las señales llevadas por una frecuencia variable. Nuestro aparato actúa al reves: una corriente eléctrica de frecuencia fija ayuda a detectar señales producidas por el débil movimiento en un condensador diferencial del péndulo de mercurio.

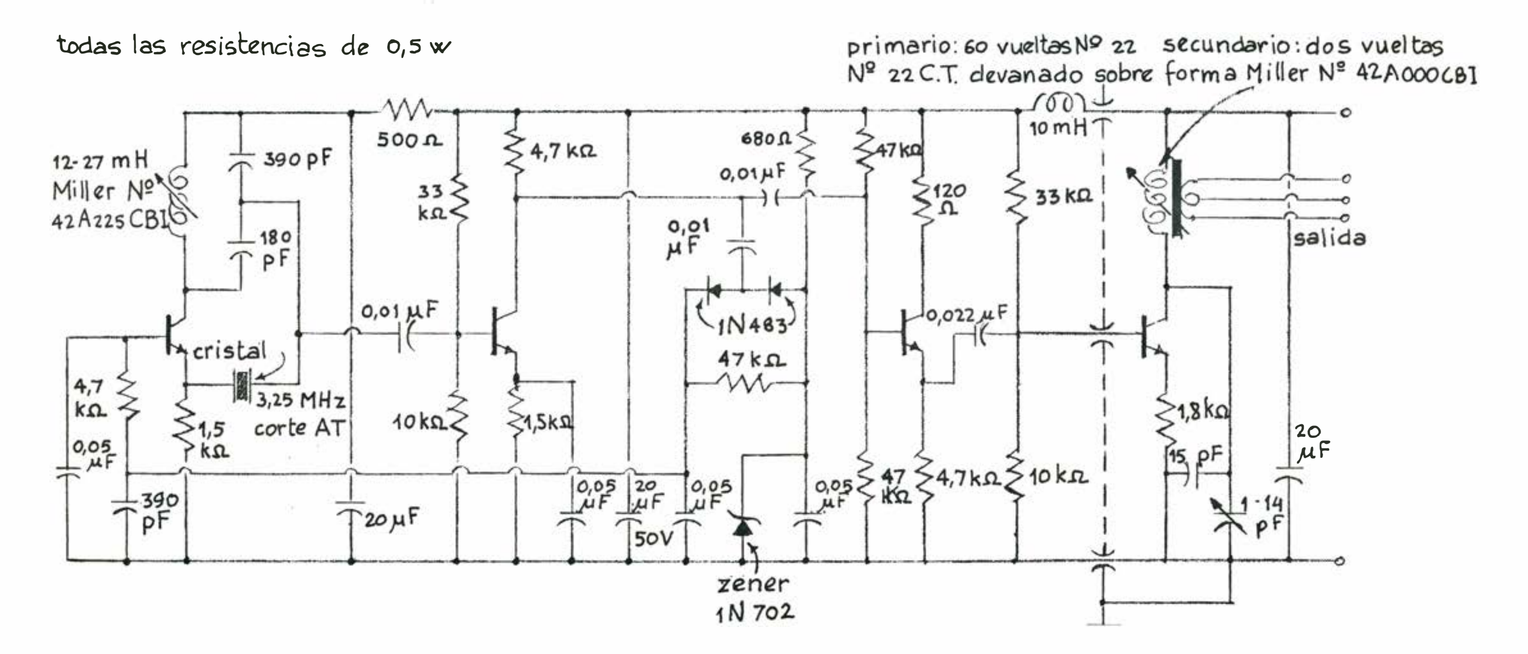
"Los circuitos discriminadores son, fundamentalmente, sistemas resonantes. Son análogos a péndulos, a las cuerdas de los instrumentos musicales y demás mecanismos similares que vibran con la máxima amplitud cuando son excitados por un estímulo de período igual al propio del sistema. El principio básico se aclara con un circuito eléctrico formado por un condensador y una inductancia, que es resonante a una determinada frecuencia. Cuando se depositan cierta cantidad de electrones en una placa del condensador, pasan a través de la inductancia para acumularse en la otra placa, vuelven a la primera placa en la fase siguiente, y así sucesivamente, con un movimiento parecido al del péndulo. Cuando el sistema recibe un impulso, vibra de un modo continuo con un ritmo característico.

"El inductor de tales circuitos generalmente consiste en un pequeño carrete con alambre aislado. Supondremos que el condensador y el inductor están conectados en serie con un generador eléctrico que funciona con una frecuencia más baja que la frecuencia propia del circuito, y que la frecuencia del generador va aumentando gradualmente hasta que supera a la propia del circuito resonante. Supóngase que se dispone de un medio para medir la diferencia de potencial entre los bornes del carrete. Se observará que la tensión en el carrete es mínima cuando la frecuencia impuesta por el generador es más baja que la frecuencia propia del circuito, que la tensión es máxima cuando hay resonancia y que disminuye a medida que la frecuencia del generador supera la frecuencia propia del circuito.

"La tensión alterna en el carrete puede convertirse en continua intercalando un diodo en el circuito; pueden amortiguarse sus oscilaciones conectando en el circuito un segundo condensador que actúa como filtro. El conjunto constituye un detector de rampa. Se ha empleado en los aparatos de radio y televisión de poco precio para detectar las señales en frecuencia modulada. En tales aplicaciones el detector de rampa está sintonizado de modo que la frecuencia media de la emisora esté localizada hacia la mitad de la característica que queda por encima de la resonancia, tal como está indicado por la letra *X* en la gráfica del detector. Cuando la frecuencia impuesta se separa de dicho punto las tensiones de salida del detector de rampa se separan de modo proporcional. La tensión y la linealidad obtenibles a la salida pueden mejorarse conectando, uno tras otro, dos circuitos.

"Nuestro discriminador inverso funciona al revés. Una corriente alterna de frecuencia fija es impuesta a través de un circuito de resonancia variable. La señal procede del movimiento del condensador variable diferencial del péndulo de mercurio. Sintonizamos los transductores al 50 por ciento del pico de salida resonante. Para un potencial de entrada de un voltio entre pico y pico, el pico de salida de cada transductor es de unos 30 voltios. Las salidas de los transductores son detectadas, filtradas y combinadas mediante tres resistencias, con objeto de suministrar una salida final única. En la práctica cada circuito transductor está conectado en paralelo con un condensador de ajuste (trimmer) del tipo de pistón. Con estos trimmer pueden compensarse pequeñas variaciones en las características de los carretes y capacitancias.

"La frecuencia constante es producida por un oscilador gobernado por un cristal. En el circuito de éste puede emplearse un transistor de silicio con una tensión colector-emisor de 45 voltios y una disipación de 0,3 vatios. El colector debería regularse a 0,05 o más amperios, y la pequeña señal, neutro-emisor-razón de transferencia *hfe*, debería ser próxima a dos a una frecuencia de 30 megahertzios. Incluso en los catálogos desechados se encuentran listas con muchos transistores que cumplen o exceden de estas especificaciones. Los experimentadores no habituados a las técnicas de los circuitos resonantes y los osciladores a base de transistores convendrá que



Esquema del circuito del oscilador de cristal de 3,25 megahertzios

busquen la ayuda de un radioaficionado.

"Una vez completo el instrumento conviene instalarlo sobre cimientos firmes en un ambiente de temperatura lo más constante posible. Nuestros sismógrafos e inclinómetros están en túneles que se encuentran a varias decenas de metros bajo tierra, y están equipados con servomecanismos que compensan automáticamente los cambios de temperatura. Sin embargo, un instrumento de características excelentes puede instalarse simplemente en un sótano o semisótano, principalmente si se aísla mediante una caja de paredes gruesas de styrofoam. Hay que usar mercurio destilado. Se necesitan unos 200 gramos para un instrumento de 60 cm de longitud.

"Se eleva unos 30 centímetros uno de los extremos del instrumento para eliminar en lo posible las burbujas de aire del mercurio y del tubo de conexión. Acto seguido se baja rápidamente. Se ajusta la base de modo que quede horizontal mediante un nivel de burbuja. Se reduce al mínimo la capacidad de los trimmer de los transductores. Se establece una unión entre el punto de enlace de las tres resistencias de un megaohmio y el punto neutro.

"Se da paso a la alimentación del oscilador. Mediante el tornillo de ajuste se reduce la distancia entre la superficie del mercurio y la placa censora de uno de los recipientes hasta que el aparato señale medio microamperio. De modo semejante se ajusta el otro recipiente.

"Se ajusta la capacidad de los trimmer hasta que cada aparato marque 3,5 microamperios. Se sustituye la unión indicada por un voltímetro de alta impedancia, el cual deberá señalar cero voltios. Si no ocurre así, se modifica la posición de una de las placas sensoras con el mercurio hasta que el voltímetro señale cero voltios. Con esto el instrumento queda equilibrado.

"Finalmente, se separa hasta el tope una de las placas sensoras. El voltímetro señalará un potencial de 17,36 milivoltios. Si marca más, la sensibilidad del aparato es excesiva. Auméntese la distancia entre las placas y el mercurio en ambos recipientes. Acto seguido, se incrementa la capacidad de los trimmer hasta que el instrumento señale de nuevo 3,5 microamperios. Si la tensión de salida es menor que 17,36 milivoltios se procede al revés puesto que la sensibilidad no es adecuada. Si se desea emplear el aparato exclusivamente como sismómetro, y con ello se pretende la mayor sensibilidad posible, se ajustan los trimmer de modo que tengan la menor capacidad y los condensadores de los recipientes posean la capacidad máxima.

Libros

Compilación de las obras de Merton y un método elemental de telecomunicación

L. Alonso, A. Comella y Ph. Morrison

LA SOCIOLOGIA DE LA CIENCIA, Robert K. Merton: versión española de Néstor Alberto Míguez. Alianza Universidad, Madrid, 1977. Sólo cuando los productos de la ciencia y sus aplicaciones sacuden nuestra rutina, hasta el punto de poner en peligro la pervivencia en el planeta, nos preguntamos por ella, la sentimos como algo amenazadoramente nuestro. Por ejemplo, la bomba de neutrones, que acaba con los tejidos humanos pero no con la textura de los edificios, y la posibilidad más o menos fundada de crear virus antes de disponer de los antivirus específicos mediante las técnicas avanzadas de la recombinación del ADN, flotan en el ambiente y han disfrutado estos días de media página en los periódicos. Hay también una confusa esperanza de que la ciencia sea el recurso *in extremis* de los males que azotan al mundo: hambre, polución, falta de recursos energéticos y de materias primas, etcétera. Sin embargo, en cuanto nos despegamos por unos instantes de las cosas inmediatas y tomamos perspectiva, observamos que la ciencia está inmersa en la sociedad contemporánea, a todos los niveles, como lo está el agua en la argamasa hecha con cemento. La ciencia da dureza y consistencia a la sociedad y a los diversos sectores que la componen, a la industria, a la agricultura y a los servicios.

Esa fusión unas veces es suave y otras violenta. Suave fue la introducción de la geometría por los agrimensores, la incorporación de todos los descubrimientos (imprenta, brújula, fundición del hierro, la cartografía cuantitativa y los transistores). Violentos fueron los choques cuando el avance científico se enfrentaba con los supuestos religiosos, esotéricos o políticos. Aristóteles, Copérnico, Galileo, Lamarck, Darwin y Lysenko ejemplifican una toma de partido científico no exento de concomitancias marginales a la empeiría. Hasta el punto de que sus conclusiones no se aceptaban como verdaderas o falsas sino como ortodoxas o heterodoxas, no viene al caso con respecto a qué credo.

Estas relaciones evidentes entre ciencia y sociedad se vienen estudiando desde los tiempos de los propios fundadores de la sociología, y aun antes. Francis Bacon (1561-1626) sostenía que la investigación filosófica, saber supremo, requiere un previo examen de los prejuicios (ídolos) que pueden ocultar la verdad, prejuicios extendidos entre los individuos y en la sociedad que componen. Auguste Comte, sus antagonistas Proudhon y Marx (rivales éstos entre sí, a su vez) y el predecesor común a todos, Saint-Simon, al poner los fundamentos de las ciencias sociales, hacen hincapié en ese punto. Se siguió planteando de alguna manera en la generación siguiente: Durkheim y Lévy-Bruhl; aunque las interrelaciones que conectan a la sociología con el conocimiento científico fueran interpretadas de forma distinta. Marx, que enlaza en su exposición con el positivismo comtiano, sentará las bases de la teoría de las ideologías como adscritas a estructuras subyacentes de tipo económico, una variante de lo cual sería la exposición reciente del saber como pura "arqueología" (es decir, hay que buscar el sentido de las palabras en el tiempo en que nacieron y dentro de los presupuestos intelectuales en que se inscribieron). El terreno estaba abonado para cuando llegaran los grandes teóricos de la sociología del conocimiento: Max Scheler, Karl Mannheim y Pitirim Sorokin.

Merton recibe esa herencia europea, de corte especulativo, y la contrasta con la sociología americana, más proclive a atenerse a las cosas tal como están y suceden en su momento: es decir, prefiere exponer quiénes son los intelectuales, los científicos, los técnicos, cómo se mueven, cómo los acepta y aúpa la sociedad, cómo se relacionan entre sí y con el resto de sus conciudadanos. Saint-Simon generalizaba, sin miedo a error: "dans tous les temps et chez tous les peuples, on trouve entre les institutions sociales et les idées une correspondance constante" (*Industrie*, vol. III, 1818). Ni siquiera el más "sociólogo" de los predecesores logra librarse de esa tendencia teorizante; así, para Durkheim, lo que importa es que "même des idées aussi abstraites que celles du temps et de l'espace sont, à chaque moment de leur histoire, en relation intime avec la structure sociale correspondante" (*Année sociologique*, 1901-1902). Y no digamos nada cuando la sociología del conocimiento se hace disciplina autónoma dentro del campo de la sociología, tal vez la más especulativa de todas: "Este libro se ocupa del problema de cómo los hombres piensan realmente. La finalidad de estos estudios es la de investigar, no cómo el pensar aparece en los textos de lógica, sino cómo funciona realmente en la vida pública y política en cuanto instrumento de acción colectiva" (Karl Mannheim, *Ideología y utopía*, Madrid, 1966). Mannheim acaba de aludir a otra corriente, positivista, de la teoría de la ciencia como análisis lógico de la fundamentación del conocimiento, cuyos principales eslabones habrían de ser Popper, Kuhn y Lakatos.

Merton centra su atención en la manera en que la ciencia, como sistema de relaciones sociales, opera dentro de la sociedad; insiste, pues, en las reglas y normas sociales de la ciencia, y en las recompensas de éstas para un mejor funcionamiento. A ese enfoque se le ha llamado "estructural-funcionalista". Cuatro son las famosas normas institucionales que describe Merton: universalismo, "comunismo", desinterés y escepticismo organizado. *"El universalismo* halla inmediata expresión en el canon de que las pretensiones a la verdad, cualquiera sea su fuente, deben ser sometidas a *criterios impersonales preestablecidos:* la consonancia con la observación y con el conocimiento anteriormente confirmado". Es el desarrollo argumentado de la famosa frase de Pasteur: el sabio tiene una patria, la ciencia no. *"El 'comunismo',* en el sentido no técnico y extendido de propiedad común de bienes, es un segundo elemento integrante del *ethos* científico. Los hallazgos de la ciencia son un producto de la colaboración social y son asignados a la comunidad". Ni siquiera la ley que lleva el nombre de su formulador pertenece a éste una vez ha pasado a engrosar el tesoro de la ciencia universalista. "La ciencia, como en el caso de las profesiones en general, incluye el *desinterés* como elemento institucional básico. El desinterés no debe ser confundido con el altruismo, ni la acción interesada con el egoísmo." El *ethos* científico que Merton esboza implica obviamente un *ethos* moral a salvo, como regla común, de motivaciones espurias. Por último, "*el escepticismo organizado* está relacionado de varias maneras con los otros elementos de *ethos* científico.

Es un mandato metodológico e institucional. La suspensión temporaria del juicio y el examen independiente de las creencias en términos de criterios empíricos y lógicos periódicamente han envuelto a la ciencia en conflictos con otras instituciones".

Estas normas, que vertebran y dan cuerpo a la sociología de la ciencia según Merton, constituyen el blanco de sus opositores. Así, T. S. Kuhn arremetía en un escrito célebre contra las mismas porque ni teórica ni empíricamente le parecían aceptables (*The Structure of Scientific Revolutions,* Chicago, 1962). Pecan de academicistas, dicen los antimertonianos, en un contexto en que los científicos son parte integrante del mecanismo y de la dialéctica industrial. Y en parte no les falta a éstos razón. La proletarización del intelectual es un hecho irrefutable en todos los países. Lo que comporta de inmediato un proceso de "tecnificación" y de "burocratización" del científico. La solución parece andar por la vía media: los técnicos, al contrastar los principios con el rendimiento industrial, al ensayar nuevos métodos, pueden aportar *in vivo* las dificultades que el científico nunca vería desde su glorieta de Academos.

Un signo del escaso interés por esta disciplina es el eco apagado que las obras de los constructores de la misma han tenido en nuestros países. Existen traducciones de los clásicos de la sociología, de Mannheim, de Barber (uno de los colaboradores más asiduos de Merton), de Popper..., pero, si exceptuamos a Mario Bunge, no existe entre nosotros una investigación coherente sobre el problema de la sociología y la filosofía de la ciencia. (L. A.)

MEDICIONES EN INGENIERIA C. V. Collett - A. D. Hope Editorial Gustavo Gili S. A. - 1976, 327 páginas, 305 figuras. No es frecuente tener la satisfacción de hallar una obra como la de los profesores Collett y Hope que, además de su interesante contenido, haya sido trasladada fielmente a nuestro idioma con tanta pulcritud y conocimiento del tema. La obra consta de 11 capítulos que abarcan los siguientes temas: I. Características de los sistemas de medición; II. Interferometría; III. Medición de dimensiones y ángulos; IV.Valoración del acabado superficial; V. Medición de la deformación; VI. Medición del tiempo, la velocidad, la aceleración y la frecuencia. VII. Medición de la fuerza, el par y la potencia; VIII. Medición de la presión; IX. Medición del caudal de los fluidos; X. Viscometría; XI. Medición de la temperatura.

De su contenido se desprende que, aparte los temas generales, los autores han prestado especial atención a las mediciones que, en el campo de la mecánica, puedan interesar a los ingenieros. Ello es natural pues sus autores son profesores de ingeniería mecánica en el Politécnico de Birmingham.

En los distintos capítulos se exponen y razonan con precisión los métodos que pueden seguirse en cada caso, estudiando su fundamento, su ejecución en la práctica y el margen de error que cabe esperar. El desbordante incremento en el número de mediciones que es preciso efectuar en la actualidad en industrias y en laboratorios, y el progreso constante en la instrumentación, obligan a documentarse para poder elegir el procedimiento más adecuado para cada caso. En este sentido, la obra tiene un especial interés para quienes enseñen ingeniería física, y para los ingenieros y técnicos que estén al servicio de la industria. El cuidadoso respeto al sistema internacional de unidades, SI, es digno complemento a la seriedad de la obra. (A.C.)

LENGUAJES SILBADOS (WHISTLED LANGUAGES) R. Busnel y A. Classe. Springer-Verlag. Han transcurrido veinte años desde que el profesor Classe explicara en SCIENTIFIC AMERICAN el notable lenguaje llamado Silbo, con el que los pastores y agricultores de la Gomera, islas Canarias, mantienen orgullosamente un sucedáneo de los radiotransmisores civiles entre los escarpados y áridos paisajes de una isla llena de profundas quebradas. Ambos autores han escudriñado a fondo desde entonces –ambos "hablan" el Silbo de corrido– esta notable hazaña cultural, conocida por forasteros desde el siglo XVII, pero mal comprendida. Resulta que la Gomera no es la única: en el valle de Ossau, en lo alto de los Pirineos franceses, una sola aldea, Aas, elaboró un sistema análogo, que ha menguado y casi desaparecido conforme la gente abandonaba sus recias casas montañesas. En las altiplanicies del noroeste de Turquía, cerca de la costa del mar Negro, toda "una población de montañeses", unos 30.000 entre hombres, mujeres y niños, dependen del silbido tal como nosotros dependemos de los teléfonos de la ciudad. Los tres sistemas son notablemente parecidos: casi con seguridad por convergencia y no por difusión, puesto que cada uno de ellos se basa en su propio idioma, francés, castellano o turco, y no se sugiere contacto mutuo alguno. (Entre los mazatecos, pueblo situado al norte de Oaxaca, México, se usa un cuarto lenguaje silbado, de especie un tanto diferente.)

Este libro es una breve monografía que resume en pocas páginas lo que se sabe y el cómo se supo. Espectros de sonidos, fotos de paisajes, cuidadosos análisis lingüísticos y fonéticos, mediciones de intensidad en función de la distancia, pruebas experimentales de aspectos críticos e indicaciones (nada fáciles) respecto a cómo hacerlo, con fotografías de adeptos, todo ello se ofrece en una obra que al lector le duele interrumpir, aunque esté destinada a especialistas en lingüística. ¿Por qué existen el Silbo y sus congéneres? La respuesta parece clara: funcionan de manera adaptativa. Los paisajes y la vida laboral de los tres lugares se parecen. No hay caminos y los senderos son tortuosos y difíciles. Los campos cultivados están situados muy arriba, en montañas bastante estériles y, en el caso de Turquía, formando terrazas. Una población aislada ha tenido por largo tiempo motivos para pensar acerca de un canal de comunicaciones que tuviese un alcance mayor y más nítido que un grito, y con menor esfuerzo. Una vez construido, el lenguaje se convierte –como ocurre en todos los logros de la sociedad humana– en motivo de orgullo y en distintivo de la comunidad, teniendo, en estos casos, la ventaja de resultar oculto para los forasteros. (En la guerra civil española, se enviaron alguna vez mensajes en Silbo, pero había gomeranos en ambos bandos, de modo que pronto se desarrolló el conocido drama de medidas y contramedidas.)

¿Funciona? He aquí los datos, tanto en general como los tomados de mediciones directas de la propagación del sonido en condiciones naturales. Los silbadores se percatan perfectamente de las sombras acústicas, de origen topográfico y térmico. El ruido es escaso; excepto cuando soplan los alisios, la Gomera da "la impresión de profunda quietud". (De vez en cuando los mirlos dan señales falsas, ¡imitando las llamadas humanas!) Una visual despejada ayuda; la vegetación amortigua muchísimo la señal. En valles boscosos y abiertos, el alcance útil es de uno o dos kilómetros, pero en valles rocosos y angostos pueden alcanzarse de ocho a diez. Los silbadores no usan sus labios, pero poniendo la lengua sobre los dientes, o uno o dos dedos en la boca, generan tonos bastante puros y penetrantes, de unos 120 decibelios a un metro de distancia, fuertes como el sonido de una trompeta.

La información se coloca en esta portadora mediante modulación de frecuencia y también mediante los ritmos de amplitud de arranque/parada del habla. Las consonantes y pausas se codifican introduciéndolas en el tono mediante confi-

guraciones naturales de rápido cambio de altura, siguiendo muy de cerca la modulación de la voz corriente durante las consonantes, puesto que la fuente de ambas cosas son los movimientos de la boca. Cualquiera puede intentar sencillamente silbar una frase articulada, sobre todo si lo hace en un idioma abierto como el castellano. Después resulta convincente leer el estudio detallado del idioma silbado. Es una recodificación fonética casi completa del habla; los usuarios pueden reconocer y repetir palabras de un idioma desconocido, si están correctamente silbadas. Según cuidadosos ensayos directos, se reconocen palabras aisladas con un 30 por ciento de error, que puede compararse con sólo un 4 por ciento en el habla normal. Lo que hace totalmente funcional el canal es la redundancia existente en mensajes más largos.

El caso mexicano es diferente. Silban niños y hombres, pero no las mujeres adultas; se trata de silbar con los labios, con un volumen mucho menor. No se practica a distancias de más de unos centenares de metros; se usa de cerca (en el mercado, por ejemplo), a diferencia de los otros lenguajes. Entre los miembros de este pueblo el trabajo en grupo es más habitual y no hay rebaños de ovejas o de vacas distantes. La codificación es muy diferente: su lenguaje hablado es tonal, y en el lenguaje silbado se presenta solamente la línea melódica del habla. De este modo, la frase tan sólo se esboza, traspuesta a un silbido labial agudo, modulado dentro de una gama estrecha de frecuencias, de aproximadamente media octava en torno a los 200 hertzios. (La frecuencia del Silbo abarca hasta dos octavas, aunque gran parte del tiempo la portadora se mantiene cercana a una misma media, que es la región de la detección óptima por sobre el ruido natural de fondo.) Los mazatecos no pueden reconocer ninguna palabra que carezca de la estructura tonal de una palabra existente en su propio idioma; su silbido no constituye una representación fonética completa. En efecto, según parece, usan la misma forma silbada para nombres prosódicamente parecidos, tales como Modesto y Gustavo.

Un capítulo final aborda las señales animales, los lenguajes de tambor y semejantes. Quizá los delfines podrían aprender el Silbo; al igual que los canarios, pueden imitar los silbidos. Los lenguajes silbados son casi tan rápidos como el habla corriente y son mucho más eficientes que el lenguaje, muy redundante, de los tambores. He aquí una tecnología popular de telecomunicaciones, probadamente eficaz y sin nada de oculto, místico o telepático (P. M.).

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

LAS NEGOCIACIONES SALT

ANNUAL DEFENSE DEPARTMENT REPORT, FY 78. Donald Rumsfeld. Government Printing Office, enero, 17, 1977.

CRUISE MISSILES. Kosta Tsipis en *Investigación y Ciencia,* n.º 7, págs. 20-31, abril, 1977.

PRESIDENT CARTER DISCUSSES STRATEGIC ARMS LIMITATION PROPOSALS en *The Department of State Bulletin,* vol. LXXVI, n.º 1974, págs. 409-414; 25 abril, 1977.

OBJETOS BL LACERTAE

QUASI-STELLAR OBJECTS. Geoffrey Burbidge y Margaret Burbidge. W. H. Freeman and Company, 1967.

RADIO PROPERTIES OF BL LAC TYPE OBJECTS. D. R. Altschuler y J. F. C. Wardle en *Nature,* vol. 255, n.º 5506, págs. 306-310; mayo 22, 1975.

THE BL LACERTAE OBJECTS. W. A. Stein, S. L. O'Dell y P. A. Strittmatter en *Annual Review of Astronomy and Astrophysics,* vol. 14, págs. 173-195; 1976.

COMUNICACIONES OPTICAS

COMMUNICATION BY LASER. Stewart E. Miller en *Scientific American,* vol. 214, n.º 1, págs. 19-27; enero, 1966.

SPECIAL ISSUE: LIGHTWAVE COMMUNICATIONS en *Physics Today,* vol. 29, n.º 5, págs. 23-60; mayo, 1976.

LIGHTWAVE COMMUNICATIONS PASSES ITS FIRST TEST. Ira Jacobs en *Bell Laboratories Record,* vol. 54, n.º 11, págs. 290-297; diciembre, 1976.

OPTICAL TRANSMISSION OF VOICE AND DATA. Ira Jacobs y Stewart E. Miller en *IEEE Spectrum,* vol. 14, n.º 2, págs. 32-41; febrero, 1977.

FLUJO DE CALOR DESDE EL INTERIOR DE LA TIERRA

CONTINENTS ADRIFT AND CONTINENTS AGROUND: READINGS FROM *SCIENTIFIC AMERICAN.* Introducción por J. Tuzo Wilson. W. H. Freeman and Company, 1976.

AN ANALYSIS OF THE VARIATION OF OCEAN FLOOR BATHYMETRY AND HEAT FLOW WITH AGE. Barry Parsons y John G. Sclater en *Journal of Geophysical Research,* vol. 82, n.º 5, págs. 803-827; 1977.

REGIONAL GEOTHERMS AND LITHOSPERE THICKNESS. David S. Chapman y Henry N. Pollack en *Geology,* vol. 5, n.º 5, págs. 265-268; mayo, 1977.

CANGUROS

LIFE OF MARSUPIALS. H. Tyndale-Biscoe. Edward Arnold, 1973.

BIOGEOGRAPHICAL CONSIDERATIONS OF THE MARSUPIAL-PLACENTAL DICHOTOMY. Jason A. Lillegraven en *Annual Review Ecology and Systematics,* vol. 5, págs. 74-94; 1974.

THE BIOLOGY OF KANGAROOS *(MARSUPIALIA-MACROPODIDAE).* Eleanor M. Russell en *Mammal Review,* vol. 4; n.º 1 y n.º 2, págs. 1-59; marzo/júnio, 1974.

LOS VACUOLOS GASIFEROS DE LAS CIANOFICEAS

BLUE-GREEN ALGAE: FINE STRUCTURE OF THE GAS VACUOLES. C. C. Bowen y T. E. Jensen en *Science,* vol. 147, n.º 3664, págs. 1460-1462; 19 marzo, 1965.

ISOLATION AND CHEMICAL CHARACTERIZATION OF GAS-VACUOLE MEMBRANES FROM *MICROCYSTIS AERUGINOSA* KUETZ. EMEND. ELENKIN. Daniel D. Jones y Michael Jost en *Archiv für Mikrobiologie,* vol. 70, n.º 1, págs. 43-64; 1970.

STRUCTURE AND FUNCTION OF GAS VACUOLES. A. E. Walsby en *Bacteriological Reviews,* vol. 36, n.º 1, págs. 1-32; marzo, 1972.

PERSPECTIVAS EN QUIMICA MACROMOLECULAR

ON THE BIOLOGICAL ROLE OF BASIC PROTEINS IN SPERMATOZOA AND DURING SPERMIOGENESIS, J. A. Subirana en *The Biology of the Male Gamete,* ed. J. G. Duckett y P. A. Racey (suplemento n.º 1 del *Biological Journal of the Linnean Society,* vol. 7, 1975), págs. 239-244.

LA ARQUITECTURA DE LA CÉLULA, BASE DE LA VIDA, Juan A. Subirana, en *Avances del Saber,* tomo XI, Enciclopedia Labor, pág. 27, Editorial Labor, S.A., 1975.

THE DUAL STRUCTURE OF CONCENTRATED NUCLEOHISTONE GELS FROM DIFFERENT SOURCES: A QUANTITATIVE ANALYSIS OF X-RAY DIFFRACTION PATTERNS, Juan A. Subirana, Fernando Azorín, Josep Roca, Joaquín Lloveras, Remedios Llopis y Jordi Cortadas en *The Molecular Biology of the Mammalian Genetic Apparatus,* ed. P. Ts'o, Elsevier/North Holland Biomedical Press, 1977, págs. 71-92.

X-RAY DIFFRACTION STUDIES OF NUCLEOPROTAMINE STRUCTURE, P. Suau y J. A. Subirana, en *Journal of Molecular Biology,* en prensa, 1977.

"SEGUNDOS MENSAJEROS" EN EL CEREBRO

THE ROLE OF CYCLIC NUCLEOTIDES IN CENTRAL SYNAPTIC FUNCTION. Floyd E. Bloom en *Review of Physiology, Biochemistry and Pharmacology,* vol. 74, págs. 1-103; 1975.

POSSIBLE ROLE FOR CYCLIC NUCLEOTIDES AND PHOSPHORYLATED MEMBRANE PROTEINS IN POSTSYNAPTIC ACTIONS OF NEUROTRANSMITTERS. Paul Greengard en *Nature,* vol. 260, n.º 5547, págs. 101-108; marzo 11, 1976.

CYCLIC NUCLEOTIDES IN THE NERVOUS SYSTEM. John W. Daly. Plenum Press, 1977.

CYCLIC NUCLEOTIDES AND NERVOUS SYSTEM FUNCTION. James A. Nathanson en *Physiological Review,* vol. 57, págs. 157-256; 1977.

JUEGOS MATEMATICOS

THE CODEBREAKERS. David Kahn. The Macmillan Company, 1967.

A FAST MONTE-CARLO TEST FOR PRIMALITY. R. Solovay y V. Strassen en *SIAM Journal on Computation,* vol. 6, n.º 1, págs. 84-85; marzo, 1977.

TALLER Y LABORATORIO

ADVANCES IN GEOPHYSICS: Vol. II, Hugo Benioff en *Earthquake Seismographs and Associated Instruments.* Academic Press, Inc., 1955.

SCIENTIFIC AMERICAN BOOK OF PROJECTS FOR THE AMATEUR SCIENTIST C. L. Stong. Simon and Schuster, 1960.